DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum politicarum

(Doktor der Wirtschaftswissenschaft)

eingereicht an der

Wirtschaftswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Dipl. Vw. Daniel Ferdinand Neuhoff

Präsident/Präsidentin der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan/Dekanin der Wirtschaftswissenschaftlichen Fakultät:

Prof. Dr. Ulrich Kamecke

Gutachter/Gutachterin: 1. Prof. Dr. h.c. Michael C. Burda, PhD

2. Prof. Fabio Canova, PhD

Tag des Kolloquiums: 07.01.2016

Reversible Jump Markov Chain Monte Carlo

-

Some Applications to Macroeconometrics

To my wife

Contents

Chapter 1

Introduction

1.1 Introduction

The four studies of this thesis are concerned predominantly with the dynamics of macroeconomic time series, both in the context of a simple DSGE model, as well as from a pure time series modeling perspective. With the exception of chapter 3, the following chapters employ a Bayesian technique called Reversible Jump Markov Chain Monte Carlo (RJMCMC). Pioneered by Green (1995), RJMCMC enables the sampling from posterior distributions over both parameters *and* models. This facilitates easier Bayesian model determination and averaging by obtaining posterior model probabilities while simultaneously exploring the model space in an efficient manner. The two major advantages featured by RJMCMC are (1) its efficient method of exploring model spaces, as it spends less time analyzing models with lower posterior probability, and (2) the ease at which posterior model probabilities can be estimated.

The following pages provide a short introduction to RJMCMC, turning then to a brief discussion of the nexus between frequentist and Bayesian econometrics, and finishing with a discussion of the motivation for this line of research and short summaries of each chapter.

1.1.1 Reversible Jump Markov Chain Monte Carlo

The following offers a short overview of RJMCMC. For a more comprehensive introduction, please refer to Fan and Sisson (2011). An overview of trans-dimensional Markov chain techniques can be found e.g. in Sisson (2005).

As noted before, RJMCMC provides a method to sample from posterior distributions spanning parameters *and* models. This is in contrast to more commonly used methods which are designed to explore posterior distributions over parameters spaces associated with one particular model.[1] The purpose of obtaining samples from the posterior distribution that also span the model space is to facilitate Bayesian model averaging and selection.[2] With fixed-

[1] Formally, let Σ_k be the parameter space associated with some model k. Fixed-dimensional samplers can only explore Σ_k for k fixed while RJMCMC enables sampling from the countable union Σ of subspaces Σ_k, that is $\Sigma = \bigcup_{k \in \mathcal{K}} \Sigma_k \times \{k\}$. See e.g. Sisson (2005).

[2] Bayesian model averaging, a paradigm put forth by Leamer (1978), essentially accounts for model av-

dimensional approaches, each model in the model space has to be considered by sampling from the corresponding posterior distribution of the parameters, computing the marginal likelihood, and then using Bayes factors and/or posterior odds ratios to compare models. In general, the number of competing models may be very large and the exploration of every posterior distribution may be computationally expensive. RJMCMC allows the researcher to explore the model space more efficiently, since the Markov chain will spend less time exploring posterior distributions for models with low posterior probability.

RJMCMC is a generalization of the venerable Metropolis-Hastings (MH) algorithm, developed in Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) and Hastings (1970), which is among the most popular algorithms for sampling from non-tractable posterior distributions by constructing a Markov chain that has a stationary or invariant distribution equal to the posterior distribution of interest. This is achieved by an accept-reject algorithm in which a new state for the Markov chain is being proposed and then accepted with some appropriately derived probability.[3] While the parameter space for a specific model is of fixed dimensionality, the dimensionality of the parameter vector may change between models when sampling across models as well as parameters. Imagine, for example, a sampler for an autoregressive time series model. When changing the number of lags to be incorporated into the model, the number of parameters varies. This situation cannot be handled by the MH approach, as will become clear in a moment.

RJMCMC provides a solution to the problem of sampling from these more involved posterior distributions. When deriving the appropriate acceptance probability for an MH sampler, the

eraging by providing an estimate of model implications weighted by posterior model probabilities. For an overview of Bayesian model averaging, see Hoeting, Madigan, Raftery, and Volinsky (1999) who also document an improved out-of-sample forecasting performance using Bayesian model averaging, which is also found by Madigan and Raftery (1994) in the context of graphical models. Kass and Raftery (1995) provide a discussion of Bayesian model selection and averaging. A recent application of RJMCMC to instrumental variable regression is presented by Koop, Leon-Gonzalez, and Strachan (2012) and Raftery, Madigan, and Hoeting (1997) discuss the merits of BMA in the context of linear regression models.

[3]Another widely used sampler is the so-called Gibbs sampler which is a special case of the MH sampler in which proposals are generated directly from full conditional distributions for each parameter given all other parameters and the data such that the acceptance probability in the accept-reject scheme is always equal to one. For an overview of both MH and Gibbs sampling see e.g. Tierney (1994).

starting point is the so called reversibility condition, or detailed balance condition, which is a sufficient condition for the Markov chain to possess the desired invariant distribution. By plugging the expression for the transition Kernel into the reversibility condition, choosing the largest possible acceptance probability following Peskun (1973), and then employing simple algebra, the appropriate acceptance probability is easily derived. This approach will, however, fail in a situation where the two sides of the detailed balance condition are not of the same dimensionality, precluding the derivation of the acceptance probability in this manner.

Green (1995) solves this problem by introducing an auxiliary proposal variable. This auxiliary variable is then combined with a differentiable bijection taking the auxiliary variable and the current state as input, such that the dimensionality of the integrals on both sides of the detailed balance condition is equalized. The appropriate acceptance probability can then be derived by a simple change-of-variables in the reversibility condition.

As an illustration, consider a sampler for a pure autoregressive model of the form:

$$y_t = P_1^p y_{t-1} + P_2^p y_{t-2} \dots P_p^p y_{t-p} + \epsilon_t;\ \epsilon_t \sim N(0, \sigma_\epsilon).$$

Here, y_t denotes the observation at time t, the P_i^p are the coefficients of the autoregressive lag polynomial of order p associated with the i'th lag, and the stochastic disturbance ϵ_t follows a normal distribution with mean zero and standard deviation σ_ϵ. In the following, P^p denotes the column vector of parameters of the lag polynomial of order p. The aim is now to obtain the posterior distribution π for the parameters of interest, namely the coefficients P_i^p and the standard deviation σ_ϵ.

Since MH samplers are only able to provide samples for distributions over parameter spaces of fixed dimensionality, the order of the lag polynomial p has to be chosen before running the estimation procedure. Denote by ς the state of the Markov chain resulting from the MH algorithm containing the values of all the parameters of interest associated with this state, i.e. $\varsigma = (P^p, \sigma_\epsilon)$. The researcher has to define prior distributions ρ for the parameters in ς, optimally representing knowledge or beliefs about the parameter values before observing the data at hand. Prior distributions are often chosen in practice to be either as non-informative as possible, or for computational or mathematical convenience.[4] Furthermore, the likelihood function $\mathcal{L}(y|\varsigma)$ giving the likelihood of observing outcome $y = (y_1, \ldots, y_t)$ given the model and parameter values has to be derived, which necessitates some distributional assumption.

Given the current state of the Markov chain ς, some new state ς' is proposed by drawing from some proposal distribution $\gamma(\varsigma'|\varsigma)$ to be chosen by the researcher guided by convenience and efficiency considerations. This proposal is then accepted with some appropriately chosen probability $\alpha(\varsigma, \varsigma')$, or correspondingly discarded with probability $1 - \alpha(\varsigma, \varsigma')$. This procedure is repeated until the desired number of samples from the posterior is obtained. The algorithm for a standard MH sampler to obtain N samples from the posterior can be summarized as follows:

Metropolis Hastings Algorithm

1. Set the initial state ς_0 of the Markov Chain
2. For $i = 1$ to N
   (a) Set $\varsigma = \varsigma_{i-1}$
   (b) Propose a new state from $\gamma(\varsigma'|\varsigma)$
   (c) Accept draw with probability

$$\alpha(\varsigma, \varsigma') = \min(1, \chi)$$

   with

$$\chi = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \times \underbrace{\frac{\gamma(\varsigma|\varsigma')}{\gamma(\varsigma'|\varsigma)}}_{\text{Proposal Ratio}}$$

   (d) If the draw is accepted, set $\varsigma_i = \varsigma'$. Otherwise, set $\varsigma_i = \varsigma$

This algorithm defines a Markov Chain which has the posterior distribution π of the parameters as its stationary distribution. As mentioned in the foregoing, the derivation of the

[4]The latter approach is philosophically not completely in accordance with the Bayesian paradigm, according to which the prior distributions should reflect prior knowledge or beliefs.

acceptance probability $\alpha(\varsigma, \varsigma')$ utilizes the reversibility condition. This condition holds if[5]

$$\int_{\mathcal{A}_\varsigma} \pi(\varsigma)\, K(\varsigma, \mathcal{B}_{\varsigma'})\, d\varsigma = \int_{\mathcal{B}_{\varsigma'}} \pi(\varsigma')\, K(\varsigma', \mathcal{A}_\varsigma)\, d\varsigma' \tag{1.1}$$

and has to hold for all subsets $\mathcal{A}_\varsigma$ and $\mathcal{B}_{\varsigma'}$ of the parameter space. In this equation, K denotes the transition kernel of the Markov chain defined by the MH-algorithm. K is given by

$$K(\varsigma, \mathcal{B}_{\varsigma'}) = \underbrace{\int_{\mathcal{B}_{\varsigma'}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma'}_{\text{Probability of moving to set } \mathcal{B}_{\varsigma'}} + \underbrace{\left[1 - \int_{\mathcal{B}_{\varsigma'}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma'\right] \delta_x(\varsigma)}_{\text{Probability of rejecting the move and } \varsigma \in \mathcal{B}_{\varsigma'}}$$

where $\delta_x(\varsigma) = 1$ if $\varsigma \in \mathcal{B}_{\varsigma'}$ and zero otherwise, see e.g. Chib and Greenberg (1995). Plugging in the formula for K into the detailed balance condition, and applying simple algebra, and taking the usual choice for the acceptance probability following Peskun (1973), yields the formula for $\alpha(\varsigma, \varsigma')$ given in the description of the algorithm.

A problem may arise if the posterior spans different models. In this case, the state ς then also contains some model indicator and the dimensionality of ς may vary between states. Consider again the case of a pure autoregressive model, where the state of the chain is given by $\varsigma = (P^p, \sigma_\epsilon, p)$. Adding one more lag to the model, going from an $AR(p)$ model to an $AR(p+1)$ model, increases the dimensionality of the parameter space associated with the state by one and the MH approach to deriving the acceptance probability will fail. Green (1995) solves this problem by introducing an auxiliary proposal variable u, with proposals sampled from some proposal distribution $\gamma(\varsigma, u)$, again chosen for convenience and efficiency, together with some appropriately chosen differentiable bijection $g(\varsigma, u)$ which maps the current state and the auxiliary variable to the proposal.

The dimensionality of the auxiliary variable u is chosen such that the dimensionality of the state vector plus the dimensionality of u is the same on both sides of the reversibility condition. This is known as the crucial dimension matching condition. Furthermore, the choice of $g(\varsigma, u)$

[5]See also Waagepetersen and Sorensen (2001).

as a differentiable bijection ensures that $d\varsigma du = |g'(\varsigma, u)| d\varsigma' du'$ where $|g'(\varsigma, u)|$ denotes the absolute value of the determinant of the Jacobian matrix of $g(\varsigma, u)$. A simple change-of-variables in the reversibility condition then allows the researcher to derive the appropriate acceptance probability.[6]

In the applications presented in this thesis, $g(\varsigma, u)$ is an identity matrix sized as shown below for a sampler for autoregressive models. In the description of the algorithm, $g(\varsigma, u)$ is denoted as $g_{pp'}(P^p, u)$ in order to emphasize the dependence on the current parameters as well as the current and proposed model orders. The determinant of the Jacobian of this mapping is equal to one. This yields the trans-dimensional equivalent of a random walk MH sampler. Proposals for the autoregressive polynomials are constructed as follows. Proposals for the moving average coefficients are obtained analogously.

Proposal Construction

1. Propose a visit to the model with order p' from $\gamma(p'|p)$
2. Draw a vector u with dimensionality $p' \times 1$ from $\gamma_{pp'}(P^p, u)$
3. Map the proposal u to the new state using $g_{pp'}(P^p, u)$:

$$\begin{bmatrix} P^{p'} \\ u' \end{bmatrix} = g_{pp'}(P^p, u) = \begin{bmatrix} A(p, p')_{p' \times p} & I_{p' \times p'} \\ I_{p \times p} & 0_{p \times p'} \end{bmatrix} \begin{bmatrix} P^p \\ u \end{bmatrix} \tag{1.2}$$

where

$$A(p, p') = \begin{cases} \begin{bmatrix} I_{p \times p} \\ 0_{(p'-p) \times p} \end{bmatrix} & \text{if } p' > p \\ \begin{bmatrix} I_{p' \times p'} 0_{p' \times (p-p')} \end{bmatrix} & \text{if } p' < p \\ \begin{bmatrix} I_{p' \times p'} \end{bmatrix} & \text{if } p' = p \end{cases} \tag{1.3}$$

The algorithm for an RJMCMC sampler for AR(p) models then reads as follows:

[6]For a detailed derivation, see e.g. Waagepetersen and Sorensen (2001).

RJMCMC Algorithm

1. Set the initial state ς_0 of the Markov Chain
2. For $i = 1$ to N
   (a) set $\varsigma = \varsigma_{i-1}$
   (b) Propose a visit to the model with order p' with probability $\gamma(p'|p)$
   (c) Draw a vector u with with dimensionality $p' \times 1$ from $\gamma_{pp'}(P^p, u)$
   (d) Set $(P^{p'}, u') = g_{pp'}(P^p, u)$
   (e) Draw σ'_ϵ from $\gamma_\sigma(\sigma'_\epsilon|\sigma_\epsilon)$
   (f) Accept draw with probability

$$\alpha = \min\left(1, \chi_{pp'}(\varsigma, \varsigma')\right)$$

with

$$\chi_{pp'}(\varsigma, \varsigma') = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \times \underbrace{\frac{\gamma(\varsigma|\varsigma')}{\gamma(\varsigma'|\varsigma)}|g'_{pp'}(\varsigma, u)|}_{\text{Proposal Ratio}}$$

   (g) If the draw is accepted set $\varsigma_i = \varsigma'$. Otherwise, set $\varsigma_i = \varsigma$

with $g_{pp'}(P^p, u)$ defined as above. The proposal $\gamma_\sigma(\sigma'_\epsilon|\sigma_\epsilon)$ for the standard deviation of the error term σ_ϵ is centered around the parameter value associated with the current state ς as in a fixed-dimensional random walk sampler. $\gamma(\varsigma'|\varsigma)$ denotes the joint probability of proposing the new state ς' given ς. For $p = p'$ this algorithm gives a standard random walk MH sampler. In practice, the major difference between the two algorithms thus lies primarily in the necessity to specify the mapping $g(\varsigma, u)$, and deriving the appropriate acceptance probability.[7]

The resulting posterior distribution can then be used for inference. For example, posterior probabilities for each model considered can be easily estimated by the share of samples for a particular model. Furthermore, by basing inference on the complete posterior yielded by RJMCMC, the researcher can easily carry out Bayesian model averaging and determination. The source codes written for the studies in this thesis are available from the author upon request.

[7] Efficient proposal distributions also have to be defined for fixed-dimensional samplers.

1.1.2 Bayesian vs. Frequentist Inference

This section gives a very brief overview of the nexus between Bayesian and Frequentist statistical methods. Most of the discussion surrounding this nexus is rather philosophical in nature and goes well beyond the scope of this thesis. A more in-depth and philosophical discussion of this topic can be found e.g. in Leamer (1978) or Freedman (1997).

While classical, or frequentist, methods provide a robust, well researched, and computationally efficient approach to statistical inference, Bayesian techniques are becoming more and more popular among researchers. This is not only due to increased research activity in this area, but also because of improved feasibility brought forth by rapid advances in computing power. Especially where the estimation of DSGE models is concerned, the Bayesian approach is by far the most popular.

The divide between Bayesians and Frequentists may seem substantial at times, starting with differences in opinion about the very nature of probability. One particularly well-known proponent of Bayesian methods in econometrics is Christopher A. Sims who titled his Hotelling lecture at Duke University in 2007 "Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere Be Bayesian". On the other hand, quite staunch "Anti-Bayesians" also exist. A lot of criticisms from this camp, e.g. as those collected in Gelman (2008), target the necessity for the elicitation of prior distributions, seen as subjective and/or arbitrary, which, in some cases, is a valid objection. Bayesians will reply to these critiques by citing theorems proving that prior beliefs or distributions become irrelevant as the amount of data grows, only to be countered in turn by Frequentists citing situations where priors indeed swamp the data.

I am in accordance with Bayarri and Berger (2004), who endorse the view that both paradigms have their advantages as well as disadvantages and form complements instead of substitutes. Used in conjunction, they enable the treatment of a question at hand from different perspectives as is done for example in the study of the dynamic properties of real per capita GDP presented in chapter 4.

1.1.3 Motivation and Outlook

This section discusses my motivation for pursuing this line of research and provides short summaries of the chapters of this thesis.

The main motivation for engaging in this area of research stemmed from the observation that assumptions generally employed for exogenous driving forces in DSGE models have very little empirical foundation and no theoretical justification. It is my view, that the structure being put on the exogenous disturbances in economic models is a way of accounting for misspecification and lack of internal propagation. It is then desirable to develop a framework that enables the researcher to model the exogenous disturbances in a data driven way. This may be done to bring the dynamics of the model closer to reality. Alternatively, the quality of a model in question may be judged by the extent to which the shock process structure influences the dynamics of the model, assuming one subscribes to the idea that the ideal model takes pure white noise as input.[8]

The second part of my motivation is based on my belief that a perfectly true model of the data cannot be found. In light of this belief it seems prudent to incorporate model uncertainty into one's analysis. Furthermore, it has been shown that incorporating model uncertainty into forecasts may, perhaps counterintuitively, improve the forecasting performance of models.[9]

To provide estimates of the structure of exogenous driving forces in DSGE models, the RJMCMC framework is particularly appealing in that it enables sampling from posterior distributions spanning not only parameters but also models, while efficiently exploring the model space. In this terminology, a DSGE model is defined by the usual collection of cross-equation restrictions *together* with the process chosen for the exogenous disturbances. Chapter 2 takes a first step towards estimating exogenous disturbances while taking the cross-equation restrictions arising from the economic model and its parameters as given. The results suggest that

[8]It may, for example, be possible to construct model selection criteria based on the spectral representation of the model, both with white noise input and the estimates of the spectrum based on the posterior, and basing model selection on the extent of the difference between the two spectra, an idea along the lines of Watson (1993).

[9]See e.g. Hoeting, Madigan, Raftery, and Volinsky (1999).

accounting for misspecification in this manner can significantly change the dynamic properties of the model compared to some arbitrarily chosen specification for the disturbances, and the results incorporate the posterior model uncertainty with respect to the exogenous driving force.

Chapter 2: Generalized Exogenous Processes in DSGE: A Bayesian Approach

This paper relaxes the usually strict assumptions placed on the structure of shock processes in DSGE models by applying RJMCMC to the estimation of an ARMA technology shock process including the order of the lag polynomials in a simple RBC model. Here, the estimation takes the calibrated model as given and the analysis focuses on the consequences of defining the process for the exogenous disturbance in a data-driven way. The impulse response functions implied by the posterior exhibit hump-shaped behavior and the original assumption of an AR(1) specification for the technology shock process is clearly rejected. Furthermore, when allowing for non-invertible moving average polynomials, a negative response of hours to a positive technology shock is contained in the credible sets.

Chapter 3: Solving and Estimating Linearized DSGE Models with VARMA Shock Processes and Filtered Data

This chapter develops a recursive solution method for linearized DSGE models with VARMA exogenous driving forces of arbitrary order. We treat the observations as a single draw from a multivariate normal distribution and calculate the autocovariances spectrally, enabling us to account for the transfer functions not only of recursive filters, but also nonrecursive filters, e.g. the Hodrick-Prescott or the Baxter-King filter, when evaluating the likelihood. This chapter has been published in Economics Letters, August 2015, Volume 133, pages 89-91.

Chapter 4: Dynamics of Real Per Capita GDP

The second to last chapter of this thesis investigates the dynamic properties of per capita GDP in six countries with particular emphasis on the persistence of the impulse responses, as measured by the sum of the coefficients of the infinite moving average representations of the estimated models. Estimates are provided from RJMCMC as well as maximum likelihood estimation with model selection according to three information criteria also used in chapter 5. Furthermore, the analysis is carried out from both a difference stationary and trend stationary perspective, as well as using the Hodrick-Prescott filter. Among the six countries studied, substantial differences in the persistence and shape of the impulse responses are documented. The countries are ranked according to the persistence estimates. This ranking is insensitive to the detrending device. Bayesian and frequentist estimates agree to a large degree. The estimates based on Hodrick-Prescott filtered data appear to be driven by filtering artifacts. For the UK, the persistence estimates are sensitive to the time period studied.

Chapter 5: Reversible Jump Markov Chain Monte Carlo vs. Frequentist Information Criteria: A Horse Race

This chapter investigates the performance of the RJMCMC framework developed in this thesis. In several experiments, the ability of RJMCMC to pick the correct model and match the true impulse response function for synthetic data is compared to the performance of three frequentist information criteria for model choice, the Akaike Information Criterion, the Corrected Akaike Information Criterion, and the Bayesian Information Criterion, together with maximum likelihood estimation for the parameters. In almost all of the experiments, RJMCMC outperforms the frequentist approaches. Where it does not, RJMCMC is at the very least competitive. Among the three frequentist criteria, the Bayesian Information Criterion delivers the best performance.

Chapter 2

Generalized Exogenous Processes in DSGE: A Bayesian Approach

Generalized Exogenous Processes in DSGE: A Bayesian Approach *

Alexander Meyer-Gohde†§ Daniel Neuhoff ‡

Abstract

We estimate ARMA (p,q) orders and parameters of the technology process in the neoclassical growth model using post war US GDP data and decisively reject the standard AR(1) assumption in favor of higher order processes. We find that the posterior concentrates density on hump-shaped impulse responses for all endogenous variables, consistent with alternative empirical estimates and the rigidities behind many richer structural models. Sampling from noninvertible MA representations, a negative response of hours to a positive technology shock is contained within the posterior credible set. While the posterior contains significant uncertainty regarding the exact order, our results are insensitive to the choice of data filter; this contrasts with our ARMA estimates of GDP itself, which vary significantly depending on the choice of HP or first difference filter. Methodologically, we use Reversible Jump Markov Chain Monte Carlo (RJMCMC) to sample from the unknown ARMA orders and associated parameters spaces of varying dimensions.

JEL classification: C11; C32; C51; C52

Keywords: Bayesian analysis; Dynamic stochastic general equilibrium model; Model evaluation; ARMA; Reversible Jump Markov Chain Monte Carlo

---

*We are grateful to Michael Burda, Fabio Canova, Marco Del Negro, and Helmut Lüktepohl as well as to participants of the 2014 CEF, 2014 SFB-649 Annual Motzen Meeting, 2014 ESEM, 2014 CFE and of research seminars at the HU Berlin and LMU Munich for useful comments, suggestions, and discussions. We are thankful for the research assistance provided by Tobias König. This research was supported by the DFG through the SFB 649 "Economic Risk." Any and all errors are entirely our own.

†Universität Hamburg, Professur für Volkswirtschaftslehre insb. Wachstum und Konjunktur, Von-Melle-Park 5, 20146 Hamburg, Germany; Tel.: +49-40-42838 3996; E-Mail: alexander.meyer-gohde@wiso.uni-hamburg.de

§Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; E-Mail: alexander.meyer-gohde@wiwi.hu-berlin.de

‡Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5667; E-Mail:daniel.neuhoff@gmail.com

2.1 Introduction

Despite recent advances in improving the fit of DSGE models to the data, misspecification remains. In his Nobel Prize Lecture, Sims (2012, p. 1202) observes that "DSGEs could be made to fit better by adding parameters allowing more dynamics in the disturbances." Likewise, Del Negro and Schorfheide (2009) identify three approaches to deal with misspecification in rational expectations models: ignore it, generalize the stochastic driving forces, or relax the cross-equation restrictions. Most analyses take the first route, ignoring this issue altogether. While Del Negro and Schorfheide (2006) follow the third route with their DSGE-VAR, the DSGE literature has not yet provided a systematic framework to address the second approach to misspecification of generalizing stochastic driving forces.[1] We fill this gap by estimating the order as well as the parameters of generalized ARMA representations of exogenous driving forces within DSGE models. Taking a Bayesian perspective, our posterior over the orders provides a quantification of model uncertainty.[2]

To accomplish the task, we adopt the Reversible Jump Markov Chain Monte Carlo (RJMCMC) methodology as pioneered by Green (1995).[3] RJMCMC provides samples from a posterior distribution spanning several, not necessarily nested, models with parameter spaces of potentially different dimensionality. In our case, each model is identified by a specific set of orders for the lag polynomials of the autoregressive and moving average components of the

[1]Some notable exceptions to the standard practice of assuming AR(1) or white noise processes for exogenous processes include Smets and Wouters (2007) who have the price-markup disturbance follow an ARMA(1,1) process, Del Negro and Schorfheide (2009) who let government expenditures follow an AR(2) instead of an AR(1) process, Alejandro Justiniano and Tambalotti (2008) who examine the robustness of the ARMA(1,1) specification for the wage and price markup shocks in the Smets and Wouters (2007) model, Croce (2014) who models long-run growth as an ARMA(1,1) process, and Cúrdia and Reis (2010) who model their vector of exogenous processes as a VAR(6).

[2]While there are certainly alternatives to our Bayesian approach, for example selecting the model with the highest maximized likelihood or using model selection criteria like the Akaike Information Criterion, we incorporate model uncertainty into the inference of statistics of interest.

[3]Markov Chain Monte Carlo (MCMC) methods have become increasingly popular for the estimation of DSGE models in recent years. See especially Fernández-Villaverde and Rubio-Ramírez (2004); along with An and Schorfheide (2007), Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010), Del Negro and Schorfheide (2011), and Guerrón-Quintana and Nason (2013) for methodological reviews and introductions; and Herbst and Schorfheide (2015) for a textbook treatment.

exogenous process, each leading to a different dimensionality of the parameter space—e.g., increasing the AR order introduces an additional parameter, increasing the dimensionality of the parameter space. The RJMCMC method rests on modifying the proposal ratios in the acceptance probability by inflating parameter vectors to common dimensionality in order to circumvent the dimensionality mismatch. We apply this approach to systematically explore the fit of DSGE models using different structures for the shock processes which, as emphasized by Brooks and Ehlers (2004), provides a computationally feasible alternative to estimating all different possible combinations of shock orders individually.[4] Again having a set of draws from the posterior allows us to quantify posterior model uncertainty and, additionally, its consequences for impulse responses and correlation structures.[5]

We begin by estimating ARMA representations of US post war GDP; we stationarize the data using two different filters, the first difference and HP filter. The posterior mode models are AR(2) and ARMA(4,5) for first differenced and HP filtered data respectively. We find that RJMCMC provides point estimates of the ARMA orders with a reliability comparable to traditional order selection criteria such as the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (AICC), and the Schwarz Criterion (SC). RJMCMC, in contrast, provides more than just point estimates by providing draws from the posterior distribution over different ARMA orders, and we find that the HP filtered GDP data is associated with substantial posterior model uncertainty with a more dispersed posterior.

We then turn to a prototypical DSGE model, Hansen's (1985) specification of the neo-

[4]The RJMCMC algorithm allows us to explore the posterior adaptively, which allows for a more efficient means of sampling across models than generating samples from the posterior of each model (for us, ARMA order combinations p and q) and then weighting according to Bayes factors. Further computational efficiency gains could be achieved, for example, with Stephens's (2000) continuous birth-death algorithm for changes in the order of processes—see Philippe (2006) for an application to autoregressive models.

[5]If multiple shocks are kept independent while generalizing their individual autocorrelation patterns, the resulting estimates admit a structural interpretation of the shocks that can guide the researcher in identifying those dimensions along which the model requires the most additional internal propagation. It may, furthermore, be possible to construct model selection criteria based on the comparison of the spectrum of variables of interest derived from estimates of the posterior with the spectrum using only pure white noise shocks giving a measure of how much structure has to be added to the model outside of economic theory, an idea along the lines of Watson (1993).

classical growth model, and relax the traditional AR(1) assumption imposed on the exogenous technology process. After confirming that RJMCMC would correctly identify the ARMA order using synthetic data generated from an AR(1) technology process, we turn to US post war GDP data and estimate the order and parameters of the technology process. We find that the data prefers higher order exogenous processes—at the mode, ARMA(3,0), but with substantial posterior density associated with other higher order specifications, such as ARMA(2,2). Strikingly, these results are not sensitive the choice of the HP versus first difference filter, with the same points estimates of order, associated parameters, and dispersion of the posterior for both. The higher order processes imply qualitatively different reactions of endogenous variables to technology shocks. Namely, the posterior impulse responses of all variables are hump shaped, reflecting common wisdom in the macroeconomics literature[6], in contrast to the monotonic responses of, say, output or labor with the traditional AR(1) process.

From a DSGE likelihood perspective, there is, without a commensurate prior specification, no reason to prefer invertible or "fundamental" representations in the presence of MA terms. Accordingly, we sample from the covariance equivalent, noninvertible representations for draws with nonzero MA order. We find a downward shift in the amplitude of the impulse responses as well as an overall increase in the posterior uncertainty regarding the impulse responses of endogenous variables to a technology shock. Strikingly, we cannot exclude the possibility of a negative response of hours to a positive technology shock, with the noninvertible MA representations closest in spirit to the news shock hypothesis.[7] Thus, we fail to reject the Gali (1999) hypothesis of a decline in hours worked in response to a technology shock, even when the impulse responses are identified by the canonical RBC model.

Our approach can be considered a Bayesian Model Averaging (BMA) method[8] for providing

[6]See especially, Cogley and Nason (1995b).

[7]See, e.g., Beaudry and Portier (2005) or Barsky and Sims (2011).

[8]For an overview see Hoeting, Madigan, Raftery, and Volinsky (1999) who also document an improved out-of-sample forecasting performance using BMA, which is also found by Madigan and Raftery (1994) in the context of graphical models. Kass and Raftery (1995) provide a discussion of Bayesian model selection and averaging. A recent application of RJMCMC to instrumental variable regression is presented by Koop, Leon-Gonzalez, and Strachan (2012) and Raftery, Madigan, and Hoeting (1997) discuss the merits of BMA

impulse responses and moments under model uncertainty, in that we weigh these statistics from different models with their respective posterior probabilities. The BMA paradigm was put forth by Leamer (1978) and interest in this approach has since increased with the advent of more powerful MCMC samplers. In a DSGE context, Wolters (2015) uses BMA to provide meta forecasts using multiple estimated DSGE models and Strachan and Van Dijk (2013) use BMA with VARs to assess the empirical support for structural breaks and the long run and equilibria restrictions implied by a prototypical DSGE model. Our analysis is close in spirit to the latter; yet, whereas they apply BMA to estimate VARs restricted commensurate with a DSGE model or to provide forecasts using various estimated DSGE models, we apply BMA to estimate the DSGE model itself.

This paper is organized as follows: We first introduce our methodology and shortly illustrate the method by constructing a sampler for a univariate autoregressive model of unknown order. Afterwards, we present the results of a small Monte Carlo study designed to gauge the power of the method for identifying univariate autoregressive moving average models using synthetic data derived from estimated ARMA models of post war US GDP data. Lastly, we apply the method to the neoclassical growth model, using synthetic AR (1) as well as post war US data, and analyze the posterior model uncertainty and its consequences for posterior impulse responses and correlations.

2.2 Reversible Jump MCMC for ARMA Processes

The Reversible Jump Markov Chain Monte Carlo (RJMCMC) methodology pioneered by Green (1995) generalizes the Metropolis-Hastings algorithm (Hastings 1970) to allow for moves between parameter spaces of varying dimensionality while maintaining detailed balance.[9] This

in the context of linear regression models.

[9] A more extensive treatment of Metropolis-Hastings samplers can be found in Chib and Greenberg (1995). See also Tierney (1998) for a comparison of RJMCMC and conventional Metropolis-Hastings kernels. Another

transdimensionality enables inference on a posterior distribution spanning several, not necessarily nested, models. In the following, we will illustrate the mechanics of RJMCMC with the construction of a sampler for univariate autoregressive models of unknown order using an RJMCMC approach.[10]

For illustration, we will derive our transdimensional random walk sampler implementation of the RJMCMC with a univariate zero mean normally distributed AR(p) model of unknown order p.[11] Such an AR(p) model is defined as

$$y_t = P_1^p y_{t-1} + P_2^p y_{t-2} + \ldots + P_p^p y_{t-p} + \epsilon_t, \;\; \epsilon_t \sim \mathcal{N}\left(0, \sigma^2\right) \tag{2.1}$$

P_i^p are the coefficients of the lag polynomial of order p associated with the i'th lag and ϵ_t is a zero mean stochastic disturbance. Denote by $P^p \doteq \left\{P_1^p, P_2^p, \ldots, P_p^p\right\}$ the vector of parameters of the AR(p) model.[12] We would like to construct a posterior distribution over the orders, p, and associated parameters, P^p, given observations on y_t.

We interpret the order of the lag polynomial p as a model indicator and will use the terms model indicator and polynomial or lag order interchangeably. The aim is now to construct a sampler for the joint posterior distribution over the different models indexed by p and their parameters. The strategy closely resembles that for Metropolis-Hastings samplers.[13] Indeed,

popular MCMC method is the Gibbs sampler which is a special case of Metropolis-Hastings samplers and ultimately RJMCMC samplers. See Gelfand and Smith (1990) for a review and comparison of Gibbs samplers as well as importance samplers and stochastic substitution and Troughton and Godsill (1998) for application to autoregressive models. Geweke (1998) provides an overview over Bayesian methods and their applications in economics. An and Schorfheide (2007) and Herbst and Schorfheide (2015), as well as Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010), Del Negro and Schorfheide (2011), and Guerrón-Quintana and Nason (2013), provide extensive treatments and introductions in the context of DSGE modeling.

[10]Several authors have applied RJMCMC to the problem of estimating univariate autoregressive (moving average) models, e.g., Brooks, Giudici, and Roberts (2003), Brooks and Ehlers (2004), and Ehlers and Brooks (2008).Relatedly, different approaches to statistical models of varying dimensionality have emerged; such as birth-death Markov Chain Monte Carlo, based on continuous time birth-death processes, as initiated by Stephens (2000) and applied to the analysis of autoregressive moving average models by Philippe (2006). A summary and comparison of these methods can be found in Cappè, Robert, and Rydèn (2003).

[11]Our derivation follows the exposition of Waagepetersen and Sorensen (2001).

[12]The part of the parameter vector associated with the standard deviation of the disturbance ϵ_t, σ will be left implicit in the exposition of this section to maintain the focus on the order, p.

[13]The appendix contains a short description of conventional Metropolis-Hastings samplers to contrast with

Metropolis-Hastings samplers are a special case in the RJMCMC framework. It is expositionally convenient to express the state of the Markov chain as

$$\varsigma = (p, P^p) \tag{2.2}$$

explicitly including the order of the autoregressive polynomial p in the state.

The detailed balance condition poses the main obstacle to the ability of transdimensional sampling to construct a joint posterior distribution over potentially nonnested models with parameter spaces of varying dimensionality. Recall the detailed balance condition (2.22),

$$\int_{\mathcal{A}} \pi(\varsigma) K(\varsigma, \mathcal{B}) \, d\varsigma = \int_{\mathcal{B}} \pi(\varsigma') K(\varsigma', \mathcal{A}) \, d\varsigma' \tag{2.3}$$

Unlike in the foregoing section, the dimension of ς can change. I.e., the state space of the Markov chain spans parameter spaces with differing dimensionality—for a sampler for AR(p) models of unknown order, when p changes so does the number of parameters. Here, the usual strategy for the derivation of the acceptance probability will fail. Green (1995) modifies the proposals in such a way that the integrals on both sides of the detailed balance condition are over spaces of the same dimensionality by introducing an auxiliary proposal variable u together with a mapping $g_{pp'}$ that maps the auxiliary proposal u and the current state of the chain to the new proposed state. The mapping $g_{pp'}$ is chosen such that the dimensionality of the integrals on both sides of the equation is inflated to some common, potentially higher, dimensionality.

In order to be able to easily verify adherence to detailed balance for a move from a state (p, P^p) to $(p', P^{p'})$, the vectors of Markov chain states and the random auxiliary proposal variables (P^p, u) and $(P^{p'}, u')$ must be of equal dimension. This dimension matching condition ensures that $\pi(P^p|p)\gamma_{pp'}(P^p, u)$ and $\pi(P^{p'}|p')\gamma_{p'p}(P^{p'}, u')$ are "joint densities on spaces of equal dimension," (Waagepetersen and Sorensen 2001, p. 54) allowing an application of a change of variables in the detailed balance equation to facilitate the construction of the tran-

the RJMCMC algorithm provided here.

sition kernel of the Markov chain. Here, $\gamma_{pp'}(P^p, u)$ is the proposal density for the auxiliary variable u going from an AR model of order p to one with order p' which may also depend on the current parameter vector P^p. The proposed new order p' is drawn from some $\gamma_p(p'|p)$ and the joint proposal density is $\gamma(\varsigma) = \gamma_{pp'}(P^p, u)\gamma_p(p'|p)$.

In our implementation of the method, we use the following differentiable bijection for $g_{pp'}$

$$\begin{bmatrix} P^{p'} \\ u' \end{bmatrix} = g_{pp'}(P^p, u) = \begin{bmatrix} A(p,p')_{p' \times p} & I_{p' \times p'} \\ I_{p \times p} & 0_{p \times p'} \end{bmatrix} \begin{bmatrix} P^p \\ u \end{bmatrix} \tag{2.4}$$

where

$$A(p,p') = \begin{cases} \begin{bmatrix} I_{p \times p} \\ 0_{(p'-p) \times p} \end{bmatrix} & \text{if } p' > p \\ \begin{bmatrix} I_{p' \times p'} 0_{p' \times (p-p')} \end{bmatrix} & \text{if } p' < p \\ \begin{bmatrix} I_{p' \times p'} \end{bmatrix} & \text{if } p' = p \end{cases} \tag{2.5}$$

This mapping leads to the transdimensional analog of a full-site updating random walk sampler. Proposals for "newly born" parameters, i.e., those $P_i^{p'}$ for $i = p+1, \dots, p'$, are centered around zero. If $p' < p$ the parameter vector is truncated and proposals for these parameters are centered around their previous values. For $p' = p$ this mapping gives a standard random walk sampler.

The detailed balance condition holds if[14]

$$\int_{\mathcal{A}_p} \pi(\varsigma)\, Q(\varsigma, \mathcal{B}_{p'})\, dP^p = \int_{\mathcal{B}_{p'}} \pi(\varsigma')\, Q(\varsigma', \mathcal{A}_p)\, dP^{p'} \tag{2.6}$$

[14]See also Waagepetersen and Sorensen (2001).

for all subsets $\mathcal{A}_p$ and $\mathcal{B}_{p'}$ of the parameter spaces associated with autoregressive polynomials of order p and p' respectively and where

$$Q\left(\varsigma, \mathcal{B}_{p'}\right) = \int_{\mathcal{B}_{p'}} \gamma(\varsigma'|p, P^p)\alpha_{pp'}(\varsigma, \varsigma')d\varsigma'$$

is the first part of the kernel in (2.24); i.e., the part of the conditional distribution of ς' associated with acceptance of the proposal.

Implementing the change of variables with the mapping $g_{pp'}$, the detailed balance condition is satisfied if

$$\pi\left(\varsigma\right)\gamma_p(p'|p)\alpha_{pp'}\gamma_{pp'}(P^p, u) = \pi\left(\varsigma'\right)\gamma_p(p|p')\alpha_{p'p}\gamma_{p'p}(g_{pp'}(P^p, u)) \tag{2.7}$$

where the details of the derivation can be found in the appendix.

Following Peskun (1973), we set the acceptance probability $\alpha_{pp'}$ as large as possible,[15]

$$\alpha_{pp'} = \min\left(1, \chi_{pp'}(\varsigma, \varsigma')\right) \tag{2.8}$$

with

$$\chi_{pp'}\left(\varsigma, \varsigma'\right) = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \times \underbrace{\frac{\gamma_p(p|p')\gamma_{p'p}(g_{pp'}(P^p, u))}{\gamma_p(p'|p)\gamma_{pp'}(P^p, u)}}_{\text{Proposal Ratio}} \tag{2.9}$$

Having chosen an appropriate acceptance probability to maintain detailed balanced, we can now implement the procedure. The resulting sequence of states approximates the joint posterior over all models indexed by their order p and the corresponding parameter vectors.

[15]Which, as noted by Green (1995), is "optimal in the sense of reducing the autocorrelation of the chain."

RJMCMC Algorithm

1. Set the initial state ς_0 of the Markov chain
2. For $i = 1$ to N
   (a) set $\varsigma = \varsigma_{i-1}$
   (b) Propose a visit to model p' with probability $\gamma_p(p'|p)$
   (c) Sample u from $\gamma_{pp'}(P^p, u)$
   (d) Set $(P', u') = g_{pp'}(P^p, u)$
   (e) Accept draw with probability

$$\alpha = \min\left(1, \chi_{pp'}(\varsigma, \varsigma')\right)$$

   $\chi_{pp'}$ is defined as in (2.9)

   (f) If the draw is accepted set $\varsigma_i = \varsigma'$. If the draw is rejected set $\varsigma_i = \varsigma$

The application to moving average models follows by analogy and the extension to autoregressive moving average (ARMA) models is straightforward. One simply defines the model indicator as a two-element vector, proposing not only visits to some model with autoregressive order p' but also for a new order for the MA-polynomial q'. Likewise, moving from scalar to vector ARMA processes of unknown order could entail choosing the maximum AR and MA orders over all the processes, in which the model indicator would remain a two-element vector and the paramters of the model would become matrices, or each process could have an individual model indication, making this now a two-n-element vector for n individual processes.

For many applications, it is desirable to restrict the parameter spaces of ARMA processes to ensure stationarity and/or invertibility.[16] To constrain sampling to these invertible and stationary regions of the parameters spaces of each model, we reparametrize the AR (and MA) polynomials in terms of their (inverse) partial autocorrelations (PACs). Details, again, are in the appendix.

[16]For the DSGE application in sections 2.4 and 2.5, we will require stationarity of the exogenous driving forces. In section 2.5, we will examine the consequences of imposing or not imposing invertibility on MA components, should they exist, on impulse responses.

2.3 RJMCMC ARMA Order and Parameter Estimation: Monte Carlo Evidence

We examine the performance of the RJMCMC method for ARMA processes of unknown order introduced in the foregoing section by carrying out two Monte Carlo experiments. For both experiments, we compare the model chosen by the posterior mode of our RJMCMC algorithm with the choices that follow from using the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (AICC), and the Schwarz Criterion (SC). We orient the Monte Carlo experiments around the same post war US per capita real GDP data[17] that will inform our DSGE model in the following section by applying our RJMCMC algorithm to obtain draws from the posterior distribution of demeaned first-differenced and HP filtered quarterly observations of US post war per capita real GDP. We find that the RJMCMC algorithm performs favorably in comparison with the three standard alternatives for identifying the ARMA orders.

2.3.1 Priors and Proposals

Table 2.1 summarizes the priors and proposals that we used in the Monte Carlo study.

Variable	Prior	Proposal
p	U(0,10)	LaplaceD(p,2)
q	U(0,10)	LaplaceD(q,2)
AR PAC	TN(0,0.25)	TN(PAC,0.0025)
MA inverse PAC	TN(0,0.25)	TN(PAC,0.0025)
σ: Standard Deviation ϵ_t	IG(1,1)	TN(σ,0.0025)

Table 2.1: Prior and proposal distributions for ARMA GDP estimation

[17] We take 1947:1-2013:3 real GDP from the NIPA tables, expressed on a per capita basis using the BLS series on the civilian noninstitutional population. Both data sets were downloaded from the St. Louis Federal Reserve's FRED database.

We choose a uniform prior over the AR and MA orders, restricting the highest allowed order to 10 for both the AR and MA polynomials. Proposals for the AR and MA orders are taken to follow a discretized Laplace distribution, LaplaceD(μ, b),[18] with location parameter, μ, and shape parameter, b, such that

$$\gamma_p(p'|p) \propto exp(-b|p-p'|) \text{ with } p', p \in [0, 1, \ldots, 10]$$
$$\gamma_q(q'|q) \propto exp(-b|q-q'|) \text{ with } q', q \in [0, 1, \ldots, 10]$$

For the (inverse) partial autocorrelations, our prior is a truncated normal distribution, denoted by TN($\mu, \sigma, -1, 1$), with location parameter, μ, and dispersion σ, and truncations at 1 and -1, imposing invertibility and stationarity. With these proposal distributions, we center the (inverse) partial autocorrelations around their previous values and new (inverse) partial autocorrelations are centered around zero.

All three standard information criteria penalize for the number of parameters in the model. This feature is also present in the posterior of our RJMCMC method with proper priors over the (inverse) partial autocorrelations. Increasing the order of, say, an autoregressive model and setting the new parameter to zero gives a model identical to the previous one with lower order; hence, does not change the likelihood. Yet, the posterior with the additional parameter is penalized as the prior probability assigned to the value of the new parameter is smaller than one, yielding a posterior probability lower than with the original, lower order. Even though the prior on the orders is uniform, the prior over the parameters behaves implicitly like an exponential prior over the sum of the orders $(p+q)$ as shown in figure (2.1).

[18] We choose the discretized Laplace over, say, a uniform over the (0,10) interval following Ehlers and Brooks (2002), who reiterate Troughton and Godsill's (1998) point that the dicretized Laplace places highest probability close to the current model, reducing the computational resources spent exploring models with a low posterior density. Consult Ehlers and Brooks (2002) for a study of different prior and proposal distributions in the context of autoregressive models.

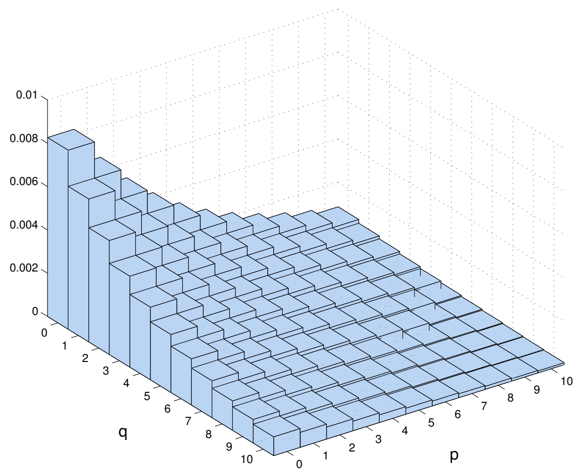

Figure 2.1: Implied prior over the orders p, q

2.3.2 ARMA Posterior Mode Estimates of US GDP

We apply the RJMCMC algorithm to obtain 3,000,000 draws from the posterior distribution of first differenced and demeaned quarterly observations of the logarithm of US per capita real GDP as well as 7,000,000 draws from the posterior distribution of the cyclical component of US GDP extracted using an HP filter with the smoothing parameter set to 1600 for the period from 1947:1 - 2013:3.[19] We employ the Kalman filter to evaluated the log likelihood, stacking higher AR and MA lags to obtain a first order vector state space.

The posterior over models using first-differenced data can be found in figure (2.2). Note that there is a substantial amount of posterior uncertainty regarding the model with textbook

[19] We generated more draws using the HP filtered data to compensate for its reduced acceptance rate. While the RJMCMC algorithm produced a total acceptance rate of 30% and an acceptance rate of 41% for proposal not involving a change in order with the first differenced filter, the numbers for the HP filtered series were 5% and 8% respectively. While Gelman, Roberts, and Gilks (1996), for example, provide conditions under which the optimal scaling in standard MCMC algorithms lies between 23% and 44%, Brooks and Ehlers (2004, p. 4) point out that "there is no Euclidean structure between models to guide proposal choices" for RJMCMC and our rates are in line with the ranges presented by Brooks, Giudici, and Roberts (2003). Stephens's (2000) birth-death RJMCMC, applied in autoregressive settings by Philippe (2006), and Fahimah Al-Awadhi and Jennison's (2004) secondary Markov chain method, along the alternatives explored by Brooks, Giudici, and Roberts (2003) provide alternatives and extensions that could potentially improve the acceptance rates in our settings.

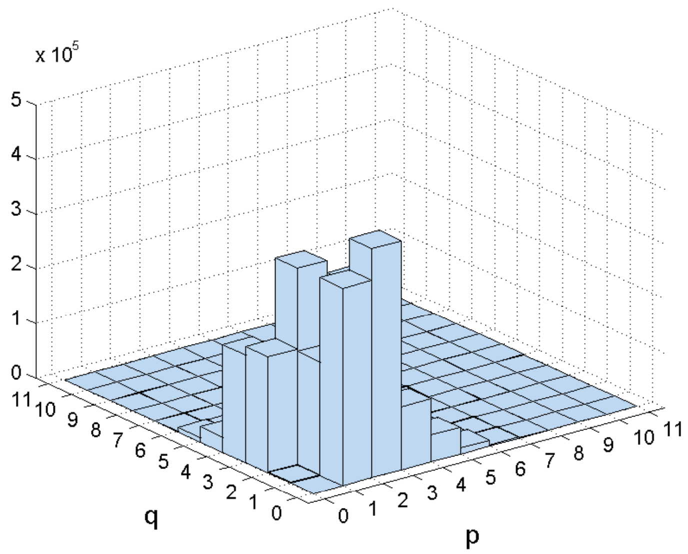

Figure 2.2: Posterior over the orders p, q for first differenced GDP data

representations such as Blanchard and Fischer's (1989, p. 9) ARMA(2,2) estimated on first differenced log GNP comfortably in the posterior distribution over models. The model at our posterior mode is an AR(2), with the posterior mean parameters conditional on this model being

$$y_t = 0.3184 y_{t-1} + 0.1297 y_{t-2} + \epsilon_t; \epsilon_t \sim N(0, 0.9025)$$

Figure (2.3) shows the posterior distribution over the orders p, q for the HP filtered data. Clearly, there is significant posterior uncertainty regarding the model reflected in the dispersion of posterior density spread over many more models than was the case with first differenced data. This is consistent with relatively high orders for the lag polynomials preferred at the posterior mode with many neighboring models mimicking the covariance structure of the model. That different filters can produce markedly different stationarized representations is well-known,[20] that the HP filtered data induces a posterior associated with higher order processes is consistent

[20]See, e.g., King and Rebelo (1993), Cogley and Nason (1995b), and Canova (1998).

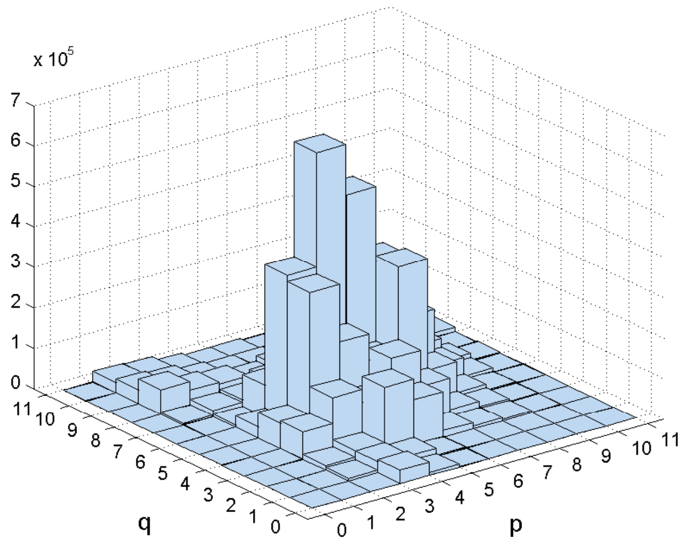

Figure 2.3: Posterior over the orders p, q for HP filtered GDP data

with the correlation functions presented by Cogley and Nason (1995b, p. 203) of HP filtered random walks. We will return to the issue of filtering when we apply the method to estimate productivity shocks inside the DSGE model in the following sections. The model at our mode is an ARMA(4,5), with the posterior mean parameters conditional on posterior mode model given by

$$
\begin{aligned}
y_t =& 0.6027y_{t-1} + 0.5304y_{t-2} + 0.0861y_{t-3} - 0.4196y_{t-4} + \dots \\
&+ \epsilon_t + 0.3786\epsilon_{t-1} - 0.2556\epsilon_{t-2} - 0.5812\epsilon_{t-3} - 0.2706\epsilon_{t-4} - 0.2154\epsilon_{t-5} \\
&\epsilon_t \sim N(0, 0.7551)
\end{aligned}
$$

2.3.3 Monte Carlo Setup

The Monte Carlo experiment is carried out by taking every 30,000th draw from the posterior for first differences and the second with every 70,000th draw from the posterior for HP filtered data, giving 100 different models each, and then for each generating 250 observations using the corresponding model and parameter values.

We implement RJMCMC by generating 1,500,000 draws from the posterior, discarding the first 1,000,000 as burn-in, and identifying the model at the mode in (p, q). The first state of the chain was set to white noise with unit standard deviation, i.e. $p = q = 0$ where p denotes the autoregressive order, q the moving average order, and $\sigma = 1$. Our metric for model choice is in accordance with a $0 - 1$ loss function, selecting the model at the mode of the posterior distribution over (p, q). It should be noted that one of the strengths of the method is the ability to quantify posterior uncertainty over models directly, such that model uncertainty can be incorporated in the calculation of posterior credible sets over impulse responses, correlations structures, or the like, providing more than just a point estimate of the model order.

We compare the model choice of our method with the choices that follow from minimizing the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (AICC), and the Schwarz Criterion (SC).[21] These are defined as

$$\mathrm{AIC} = 2k - 2\ln(\hat{\mathcal{L}}), \quad \mathrm{AICC} = AIC + \frac{2k(k+1)}{n-k-1}, \quad \mathrm{SC} = -2\ln(\hat{\mathcal{L}}) + k\ln(n)$$

with k being the number of model parameters and n the number of observations. $\hat{\mathcal{L}}$ denotes the maximized likelihood value of a model, i.e., for given ARMA orders p and q.

2.3.4 Results

We report the proportion of correctly identified models in table 2.2. The RJMCMC method outperforms the set of traditional information criteria in all cases but one. The exception is data generated from the model at the posterior mode of HP filtered data.[22]

The posterior using HP filtered data generally implies higher order processes, which hampers all methods' ability to correctly identify the model. This follows as the autocorrelation

[21] Calculations for the three standard measures were carried out using the R package auto.arima.

[22] Though an increase in the number of the draws from the posterior with our method ought to further improve its performance.

Method	First Differences	HP Filter
RJMCMC	0.23	0.05
AIC	0.08	0.03
AICC	0.09	0.02
SC	0.18	0.01

Table 2.2: Proportion of correctly identified models in the Monte Carlo experiments

structure of ARMA models of higher orders may be very close even if the orders of the lag polynomials differ and the likelihood is therefore rather flat across models. This was reflected likewise in the posterior distribution over models in the initial estimation itself, see figure (2.3). In contrast to standard metrics, however, RJMCMC enables the characterization of the resulting uncertainty regarding model selection choices and the posterior therefore provides the researcher with a tool to gauge the extent of model uncertainty.

Of course, the ability of the method to estimate the parameters of the model along with the order of the model is of importance. Figure 2.4 reports the recursive means of the parameter draws of the model parameters data conditional on the model being correctly identified (i.e., $p = 2$ and $q = 0$ for first differenced data) from one chain. These values clearly converge close to the values underlying the data generating process.

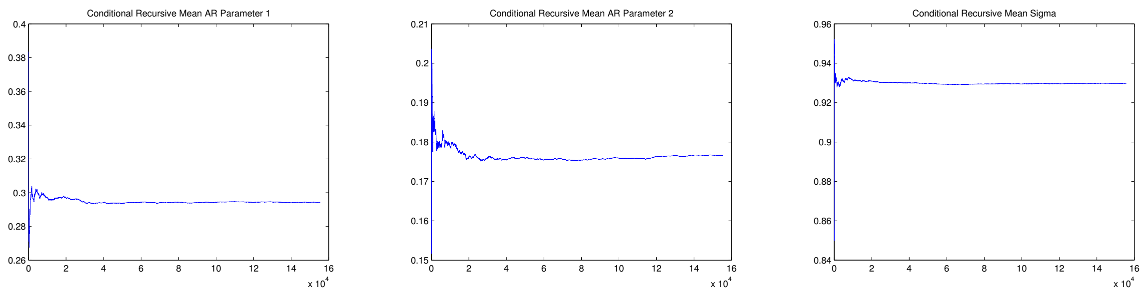

Figure 2.4: Recursive parameter means from the conditional posterior

In conclusion, the RJMCMC method exhibits roughly the same or better performance as classical methods concerning order identification while providing a complete posterior distribution over parameters and model orders that can be used for the posterior analysis of statistics of interest. With first differenced and HP filter data, US post war GDP is best represented by an AR(2) and ARMA(4,5) respectively. We are interested in posterior statistics of DSGE models such as impulse responses and correlation structures and will now turn to a DSGE setting and apply the RJMCMC method there.

2.4 Neoclassical Growth Model

To examine how the RJMCMC method can be applied to a DSGE model, we consider Hansen's (1985) specification of the neoclassical growth model. In this simple model, the social planner's problem is to maximize the discounted lifetime expected utility of a representative household given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[ln\left(c_t\right) + \psi ln\left(1 - l_t\right)\right],\ 0 < \beta < 1 \tag{2.10}$$

with c_t representing consumption and l_t hours; $\beta \in (0,1)$ is the subjective discount factor of the household and ψ weights the utility of leisure $1 - l_t$ in the household's utility function. The social planner faces the resource constraint

$$c_t + i_t = y_t \tag{2.11}$$

where investment i_t contributes to the accumulation of capital k_t through

$$k_t = \left(1 - \delta\right) k_{t-1} + i_t \tag{2.12}$$

with the depreciation rate δ and where production y_t is neoclassical and given by

$$y_t = e^{z_t} k_{t-1}^{\alpha} l_t^{1-\alpha} \tag{2.13}$$

with z_t being stationary stochastic productivity. Hansen (1985) assumed a highly autocorrelated AR(1) process—with the autoregressive parameter set to 0.95—following Kydland and Prescott (1982). Relaxing this assumption will be the focus of our investigation in what follows, so we leave it otherwise unspecified for now.

The first order conditions of the social planner's problem are given by

$$\frac{1}{c_t} = \beta E_t \left[\frac{1}{c_{t+1}} \left(1 - \delta + \alpha e^{z_{t+1}} \left(\frac{l_{t+1}}{k_t} \right)^{1-\alpha} \right) \right] \tag{2.14}$$

$$\frac{\psi}{1 - l_t} = \frac{1}{c_t} (1 - \alpha) e^{z_t} \left(\frac{k_{t-1}}{l_t} \right)^{\alpha} \tag{2.15}$$

An equilibrium is defined by the equations (2.11) through (2.15) along with a specification for the stochastic productivity process z_t.

$\overline{L}$	$\frac{1}{3}$	Steady state employment 1/3 of total time endowment
α	0.36	Capital share
δ	0.025	Depreciation rate for capital
$\overline{R}$	1.01	One percent real interest rate per quarter

Table 2.3: Model calibration

In this exercise, we will take the parameters of Hansen's (1985) calibration of all parameters outside the specification of the stochastic productivity process z_t as given. This will allow us to concentrate on the contribution of the RJMCMC algorithm in estimating the order and parameters of the exogenous process.[23] The calibrated parameters reported in table 2.3 deliver standard values for parameters, imposing , e.g., that about one third of agents' time endowment is spent in employment activities, capital contributes a little more than one third to production. As we will consider arbitrary ARMA processes for z_t, the model does not fit

[23] Ultimately, we will estimate deep parameters along with the orders and parameters of exogenous processes. We engage in the intermediate step of holding the deep parameters constant to focus on the exploration of the order and parameters of exogenous processes, avoiding the assessment of the relative influences of priors regarding orders and deep parameters on the posterior.

canonical DSGE linear problem statements, e.g., Klein (2000), which allow for straightforward calculation of the likelihood function. While we could redefine the model to include the entire state vector induced by the ARMA exogenous process as endogenous variables to bring the model into the canonical form, doing so would significantly increase the computation costs involved in the QZ decomposition for the state transition and the Sylvester equation for the impact matrix of shocks. We apply the method introduced in chapter 3 of this thesis, which solves linear DSGE models with (vector) ARMA processes of arbitrary orders directly.

2.5 Estimation Results for the Neoclassical Growth Model

We carry out three exercises using the neoclassical growth model as presented above. First, in order to check whether the method could pick up the correct underlying process for a technology shock in this model, we generated 250 observations of synthetic data using the AR(1) process as reported by Hansen (1985) in his original study. Second, we estimate the order and parameters of the technology shock process for the model using US GDP data, treated with the HP filter as in Hansen's (1985) original study. Finally, we examine the robustness of our results to this choice of filter and first difference the data instead.

2.5.1 Priors and Proposals

The priors and proposals for the shock process orders and parameters are reported in table 2.4.[24]

The priors remain the same as in the Monte Carlo study, while the dispersion parameters of the proposals were tuned using short pilot runs to increase the efficiency of the RJMCMC

[24] Our priors and proposals parallel those of the monte carlo study; see footnote 18.

Variable	Prior	Proposal
p	U(0,10)	LaplaceD(p,2.2)
q	U(0,10)	LaplaceD(q,2.2)
AR PAC	TN(0,0.25)	TN(PAC,0.0016)
MA PAC	TN(0,0.25)	TN(PAC,0.0016)
σ	IG(1,1)	TN(σ,0.0025)

Table 2.4: Priors and proposals for the RBC model estimation

algorithm.[25]

2.5.2 Synthetic AR(1) Data

For this exercise we generated 250 realizations for the technology shock according to the AR(1) specification and calibration in Hansen (1985)

$$z_t = 0.95 z_{t-1} + \epsilon_t, \ \sim (0, 0.712^2) \tag{2.16}$$

We then fed the resulting series for z_t into the linearized RBC model and applied our method to the resulting synthetic data on output, y_t. We evaluate the likelihood function recursively using the Kalman filter to generate 650,000 draws from the posterior, discarding the first 100,000 draws as burn in. Standard visual measures indicated convergence. Figure 2.5 shows the posterior distribution over the orders for the disturbance. The method places an overwhelming majority of the posterior weight on the AR(1) model—obviously correctly identifying the AR(1) data generating process for the productivity process with observations on output, y_t. Furthermore, the AR(1) parameter and standard deviation of the innovation were 0.959 and 0.762, respectively, at the posterior mean conditional on the AR(1) model

[25]As discussed in footnote 19, the usual range of desired acceptance rates does not apply in the transdimensional RJMCMC. Our acceptance rates overall and for within model moves fall in the 5% to 10% range.

having been selected.

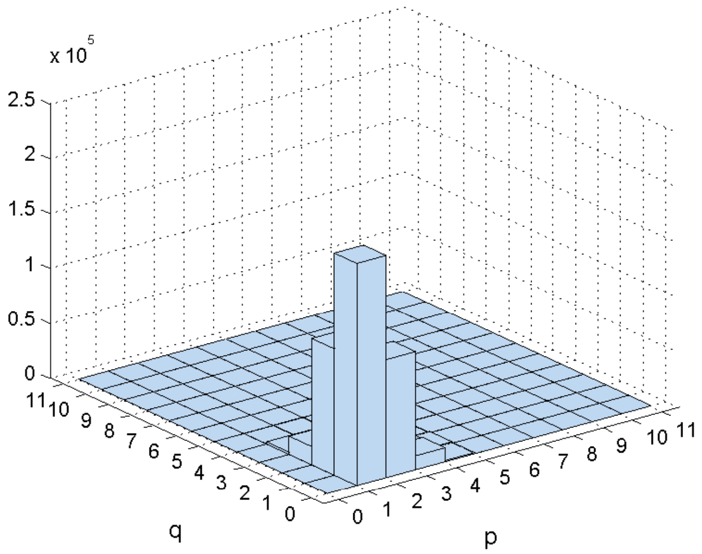

Figure 2.5: Posterior over the orders for the shock process, synthetic AR(1) data from (2.16)

This result gives us further confidence that, if the real world process for the productivity shock were AR(1), it would be correctly identified by the RJMCMC method we propose.

2.5.3 US GDP Data: Estimates

We now address what US postwar GDP data can reveal about the productivity shock in Hansen's (1985) model. We estimate the productivity shock process using HP filtered quarterly US GDP per capita as in Hansen (1985) taking his original calibration and value of 1600 for the smoothing parameter in the HP filter as given.[26] In applying the RJMCMC method introduced in section 2.2, we generated 4,000,000 draws per chain discarding the first 1,000.000 draws as burn in.[27] We apply the HP filter to the model when evaluating the likelihood,[28] thus treating

[26]See footnote 17 for details on the data series. We address robustness to the choice of filter in section 2.5.6.

[27]We initiated the five chains at $(0, 0)$, $(0, 10)$, $(10, 0)$, $(10, 10)$, and $(5, 5)$ for p and q.

[28]As the HP filter is a two-sided filter, we cannot use the Kalman filter to while applying the HP filter to evaluate the likelihood function the model. Thus, we treat the sample as one large draw from a multivariate

the data and the model with the same filter.

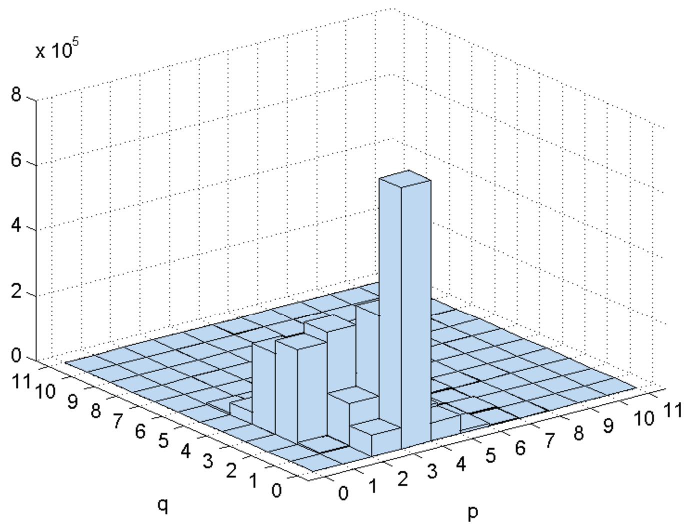

Figure 2.6: Posterior over the orders for the shock process, HP filtered data and model

Figure 2.6 shows the posterior over (p, q) for this exercise. The model at the mode of the posterior is ARMA(3,0) and the baseline AR(1) specification of Hansen (1985) is clearly rejected. There is much more substantial uncertainty regarding the correct shock process than in the Monte Carlo exercises above. The prior posterior plots in figure (2.7) indicate that our results are not solely driven by our choice of priors, likewise confirmed by comparing the posteriors over orders in figure 2.6 to the implied priors in figure 2.1.

Figure 2.8 reports recursive means of the first AR parameter for chains with differing initial states for the orders of the ARMA polynomial for the technology shock, calculated both conditional on the model at the mode of the posterior as well as unconditional means. Inspection suggests that all three chains have converged,[29] as do the the posterior statistics,

normal distribution, where we calculate the sequence of HP filtered autocovariances spectrally. See chapter 3 of this thesis for details on the computation.

[29] It is not clear, however, whether these standard graphical or other formal measures of convergence, e.g., Brooks and Gelman (1998), apply without adaptation in transdimensional analyses, see e.g., Fan and Sisson (2011).

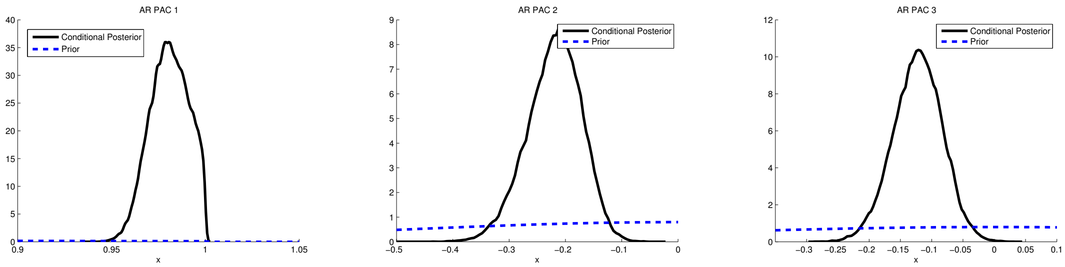

Figure 2.7: Priors and posteriors for partial autocorrelations, HP filtered data and model

such as impulse responses, that we will examine.

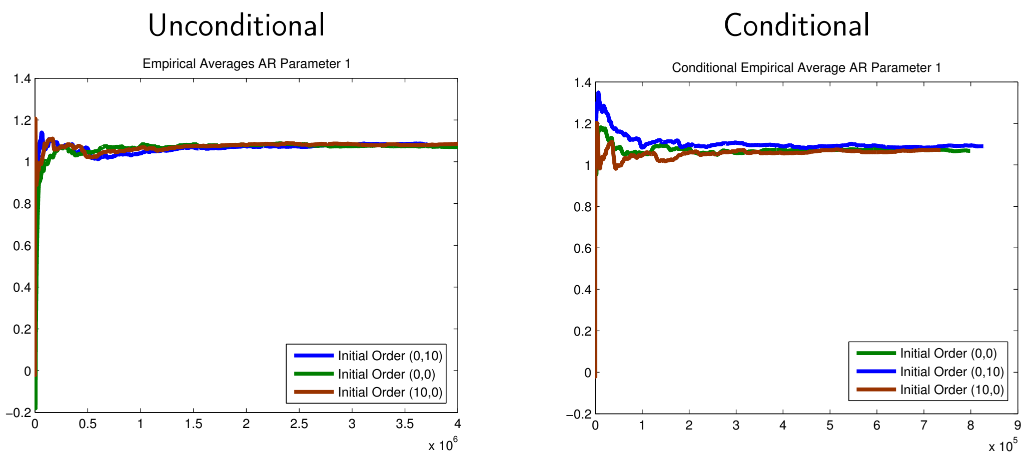

Figure 2.8: Convergence diagnostics, HP filtered data and model

Table 2.5 reports point estimates for the shock process parameters taken from the posterior distribution conditional on $(p,q) = (3,0)$; the standard deviations for the posterior means are also reported. Additionally, the first two autocorrelations of the exogenous process, z_t, implied by these point estimates are given. The first autocorrelation is higher than, though consistent with, the choice of Hansen (1985) following Kydland and Prescott (1982) to model the technology process with a near unit root.

Parameter	Mean	Median	Hansen (1985)
AR(1)	1.1689	1.1681	0.95
	(0.04)		
AR(2)	-0.0732	-0.0725	N/A
	(0.06)		
AR(3)	-0.1224	-0.1215	N/A
	(0.04)		
σ	0.5873	0.5733	0.712
	(0.08)		
$\rho(1)$	0.9804	0.9810	0.95
$\rho(2)$	0.9528	0.9542	0.9025

Table 2.5: Posterior point estimates and autocorelations, HP filtered data and model along with Hansen's (1985) original AR(1) specification

2.5.4 US GDP Data: Correlation Structure

We now examine the variance and correlation structures implied by our posteriors and compare these with the data and the statistics implied by our baseline AR(1) model implied by Hansen (1985).[30] The posterior matches the structure of the second moments of output quite well, with the role of the prior becoming relatively more important for higher order correlations. As we estimated with real per capita GDP data, this is reassuring and indicates that the procedure does indeed provide a substantial improvement in fit.

Data	Hansen	Posterior Mode Model	Posterior Mode	90% Posterior Credible Set
2.8491	3.2574	2.8332	2.8182	2.1074 — 4.0965

Table 2.6: Standard deviation of output, in %, HP filtered data and model

The standard deviations of output are listed in table 2.6. Both the standard deviation of model at the posterior mode of the ARMA order and parameter space and the posterior

[30]Following Hansen (1985), we calculate the second moments for his model using an HP filtered (with the smoothing parameter, λ, set to 1600) version of model.

mode of the standard deviations line up very close to the statistic in the data, whereas the statistic of Hansen (1985) shows greater a difference from the value in the data. The 80% posterior credible set shows the extent of posterior uncertainty, which here is great enough to encompass all the point values reported.

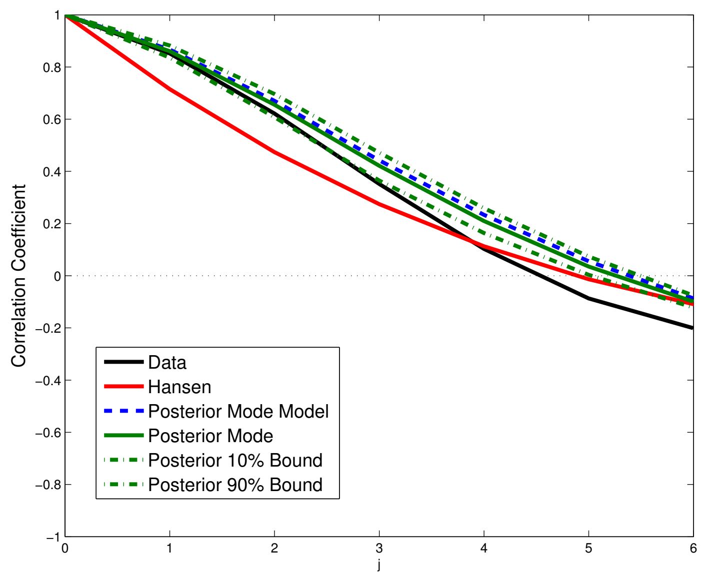

Figure 2.9: Comparison of autocorrelations of output, HP filtered data and model

The first six autocorrelations tell a more certain story, however, and can be found in figure 2.9. Again, both the autocorrelations of the model at the posterior mode of the ARMA order and parameter space and the posterior mode of the autocorrelations match the statistic in the data very closely. The AR(1) structure imposed by Hansen (1985) forces a compromise, with the initial autocorrelation being somewhat lower and the later values somewhat higher than in the data.

The fit as implied by the point estimates of our posterior with respect to our observable series output is reassuring in that our application of the RJMCMC method is successfully doing what it should. With a mean zero normally distributed process, the second moments describe the stochastic properties of the process and our posterior brings the second moments of output from the RBC model closer to the data by selecting appropriate ARMA processes.

2.5.5 US GDP Data: Impulse Responses

With a posterior distribution over both models—i.e., orders p and q—and their parameters for the ARMA technology process, we plot impulse responses taking posterior uncertainty about the model into account. In the presence of MA components, this requires us to take a stand on which covariance equivalent representation we choose.[31] We will first examine the invertible or fundamental impulse responses associated with the posterior distribution. Then, we will allow the possibility of nonfundamental representations by sampling with a noninformative prior from the admissible (i.e., real valued) covariance equivalent representations and examine the resulting impulse responses.

In figure 2.10, we plot the impulse responses to a one standard deviation technology shock. We plot the invertible impulse associated with the model at the posterior mode of the ARMA order and parameter space against the pointwise posteriors (mode and 80% credible set) over all impulse responses weighted by posterior probabilities. To guarantee invertibility, we sample from the inverse partial autocorrelations analogously to our sampling from the partial autocorrelations for the AR components that guarantees stationarity. We also include the impulse response with Hansen's (1985) AR(1) technology assumption in the plot. The data driven selection of the specification of the shock process implies a different dynamic behavior for many of the model's endogenous variables compared to Hansen's (1985) AR(1) specification. While our procedure confirms the consumption smoothing and time to build responses of consumption and capital, we identify a sluggish responses of all endogenous variables, a salient feature of the data identified in many empirical studies; e.g., Cogley and Nason (1995b), who identify a hump shaped response of output to transitory technology shocks using both an SVAR and a VEC model. In essence, the sluggishness of output in the

[31] See Lippi and Reichlin (1994), Fernández-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007), and Alessi, Barigozzi, and Capasso (2011) for more on different MA representations in macroeconomic modeling.

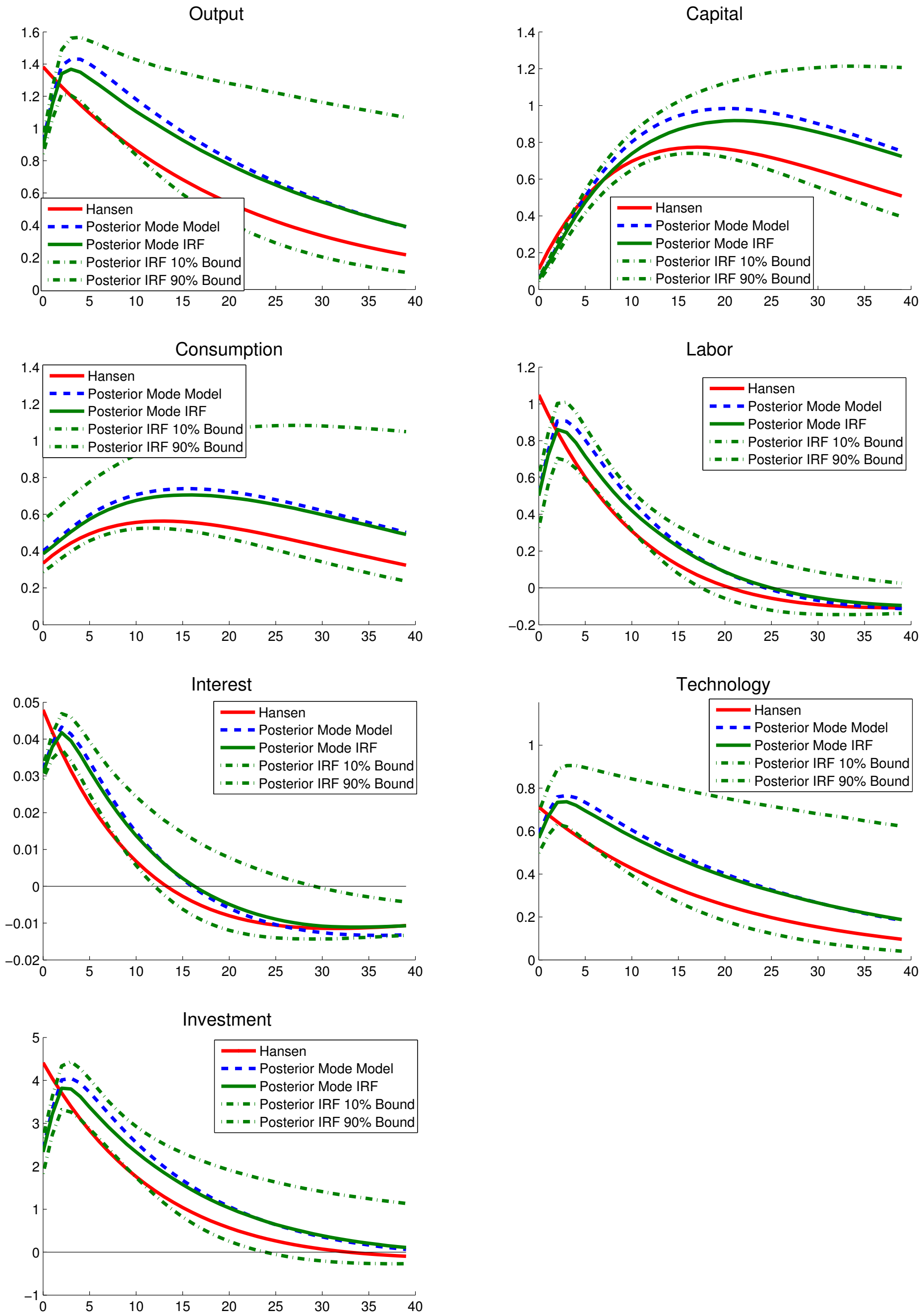

Figure 2.10: Impulse responses to a one standard deviation technology shock
Invertibility of MA components imposed
HP filtered data and model along with Hansen's (1985) original AR(1) specification

data that is captured by frictions in more sophisticated models, see especially Sims (1998) for an early assessment, is relegated to the exogenous process by our procedure.

We now move beyond imposing fundamentalness in the sampled MA components. In admitting nonfundamental MA representations, we acknowledge invertibility issues revisited in VAR contexts by, e.g., Fernández-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) that imply that the covariance structure associated with our posterior distribution potentially implies several possible different structural representations. For an invertible or fundamental moving average representation, the roots, λ_{q_i}, of the MA polynomial

$$\gamma_i(\lambda) \doteq 1 + \gamma_{i,1}L \ldots + \gamma_{i,q}L^{q_i} \tag{2.17}$$

must all be contained outside the unit circle; that is, there exists no λ such that $\gamma_i(\lambda) = 0$ where $|\lambda| \geq 1$.[32] We follow Lippi and Reichlin's (1994) root-flipping procedure and propose the following algorithm to sample from admissible covariance equivalent representations assuming uninformative priors over the different representations.

[32] See, e.g., Hamilton (1994, p. 67).

Sampling From Admissible Covariance Equivalent Representations

1. For a given draw of order $q > 0$ for the MA component of the exogenous process, factor the MA polynomial as

$$1 + \gamma_{i,1} L \ldots + \gamma_{i,q} L^{q_i} = (1 - \lambda_1 L)(1 - \lambda_2 L) \ldots (1 - \lambda_{q_i} L) \tag{2.18}$$

2. Enumerate all possible combinations of root flips. I.e., discard any combination that would flip only one of a complex conjugate pair of roots (as flipping only one of a pair of complex conjugates would lead to imaginary moving average parameters that we rule out on economic grounds), let $\tilde{n}$ denote this number of admissible combinations of root flips
3. Draw an integer $n \in \{0, 1, \ldots, \tilde{n}\}$ from a uniform distribution
4. Flip the roots according to the combination enumerated with n, where a draw of 0 indicates that no root is flipped (i.e., the invertible or fundamental representation is drawn.

For example, if $n = 10$ is drawn and the number 10 was associated with flipping roots λ_2 and λ_3, the MA polynomial for calculating impulse responses becomes

$$\gamma_i(L) = (-\lambda_2)(-\lambda_3)\left(1 - \frac{1}{\lambda_2} L\right)\left(1 - \frac{1}{\lambda_3} L\right)(1 - \lambda_1 L)(1 - \lambda_4 L) \ldots (1 - \lambda_{q_i} L) \tag{2.19}$$

Drawing the covariance equivalent representation from a uniform distribution over all admissible covariance equivalent representations puts equal weight on each admissible representation, reflecting our flat prior across the different representations over which DSGE theory is noninformative.

Figure 2.11 contains the pointwise posteriors (mode and 80% credible set) over all impulse responses weighted by posterior probabilities and drawn, potentially, from nonfundamental covariance equivalent representations as outlined above. We plot these pointwise posteriors against the invertible representation of the model at the posterior mode over ARMA orders and their parameter values, as well as against the impulse response with Hansen's (1985) AR(1) technology assumption. The admission of non-fundamental representations increases our uncertainty over the dynamic response of variables to a technology innovation, spreading the bounds of the 80% credible sets apart. Most of this spread is downward so that the number of periods for which the 80% credible set covers exclusively positive responses to a technology shock is greatly reduced.

Admitting non-fundamental moving average representations places a negative response of hours to a positive technology shock is contained in the credible set. Hence even this simplest real business cycle model with an estimated technology shock process can recreate this stylized observation of Gali (1999) and Francis and Ramey (2005). The conclusion, therefore, that the stochastic growth model is unable to generate this response to technology shocks would require a strong prior against the noninvertible moving average representations, e.g., against news and policy announcement shocks. Nonetheless, the majority of the posterior mass still lies in a region where the response of hours to technology is conventional, in line with the results in Chari, Kehoe, and McGrattan (2008) and Uhlig (2004).

In sum, the posterior mode model and the posterior distribution over impulse responses, both fundamental and admitting the possibility of non-fundamental moving average representations, as markedly different than those implied by the AR(1) assumption in Hansen's (1985) original study. The data clearly favors hump-shaped impulse responses and cannot rule out a drop in hours in response to a positive technology shock.

2.5.6 Robustness to Data Filter

We now address the sensitivity of our results to the choice of filter, which is particularly important considering the starky different results we obtained with our ARMA estimates of GDP. We choose the first difference filter as alternative to the HP filter used above, following our choice in section 2.3. This filter is one sided so we return to using the Kalman filter to evaluate the likelihood of the filtered data using the filtered model.

The posterior distribution over orders (p, q) can be found in figure 2.12. The mode model is an ARMA(3,0), just as was the case above with the HP filter. Furthermore, the entire

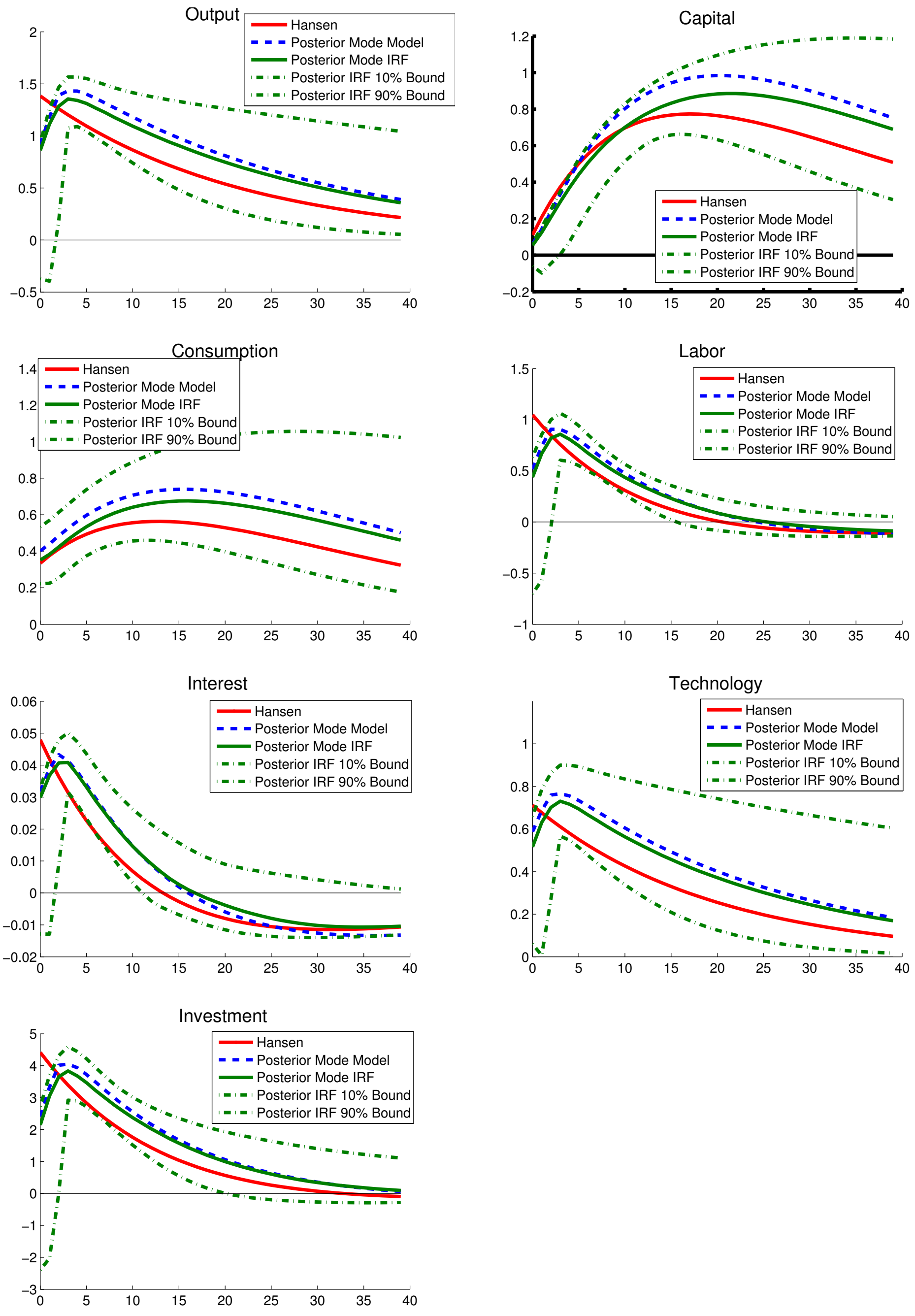

Figure 2.11: Impulse responses to a one standard deviation technology shock
Invertibility of MA components ***not*** imposed
HP filtered data and model along with Hansen's (1985) original AR(1) specification

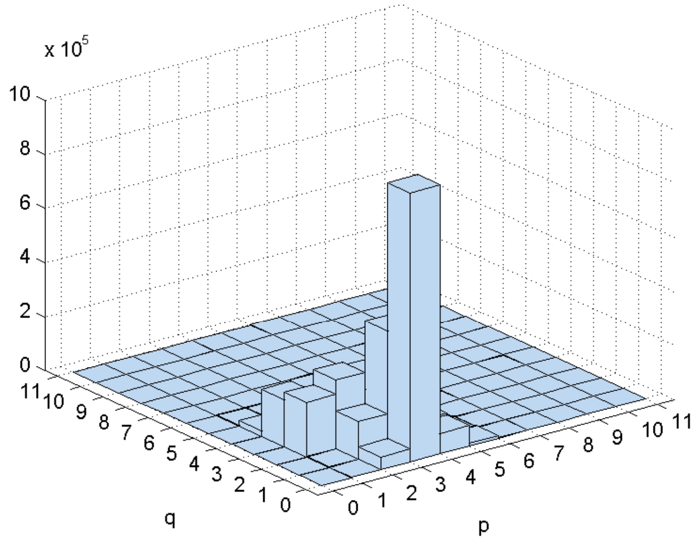

Figure 2.12: Posterior over the orders for the shock process, first differenced data and model

posterior across models in figure 2.12 is remarkably similar to the posterior using the HP filter, see figure 2.6. The posterior mean parameter estimates conditional on the ARMA(3,0) model with the first difference filter along with their standard deviations are juxtaposed with those with the HP filter in table 2.7. The extent of agreement is again striking.

We conclude that that our implementation of RJMCMC to estimate exogenous processes in DSGE models is not significantly dependent on the choice of filter. However, note that this requires that both the data and the model be treated with the same filter, otherwise we filter dependence as we found in section 2.3 would seem more likely.

2.6 Conclusion

In this paper we present a novel approach to addressing misspecification in DSGE models. Theory generally provides no guidance on the order of the exogenous processes and DSGE analyses seldom if ever estimate the order—the usual choice of the AR(1) structure on ex-

Parameter	First Differenced	HP Filtered
AR(1)	1.175	1.1689
	(0.04)	(0.04)
AR(2)	-0.0721	-0.0732
	(0,06)	(0.06)
AR(3)	-0.1289	-0.1224
	(0.04)	(0.04)
σ	0.59581	0.5873
	(0.07)	(0.08)

Table 2.7: Posterior mean estimates, first differenced vs. HP filter

ogenous processes often lacks empirical support. We relax the assumptions placed on the structure of exogenous processes and estimate generalized ARMA(p, q) processes of unknown orders. Our method treats the ARMA orders of shock processes as additional parameters to be estimated, enabling the researcher to identify those shock process structures which bring the model closer to the data.

Our ARMA estimates of US post war GDP vary with the filter chosen and we demonstrate that the RJMCMC method compares favorably with traditional methods of order identification. Turning to the estimation of ARMA process for productivity shocks in the neoclassical growth model, we find compelling evidence for higher order processes, in contrast to the standard AR(1) assumption, and find that this result is robust to the choice of HP or first difference filter. Our posterior impulse responses are markedly different than those generated under the original calibration, with the higher order processes we clearly identify hump-shaped impulse responses hump-shaped responses for all endogenous variables including output. When taking an agnostic stance regarding the invertibility of the MA polynomials in our posterior, we cannot rule out a drop in hours in response to a positive technology shock.

Our method has the advantage that it allows for the quantification of posterior model uncertainty along with the posterior parameter uncertainty in standard DSGE Bayesian estimations, ultimately enabling the analysis of a joint posterior over different specifications of the exogenous processes including their parameters as well as parameters of the model. By incorporating model uncertainty, the posterior impulse responses identified by our method offer an additional method of identifying empirical structural responses that use the entire model for identification and not just a subset of its short and/or long run restrictions. As noted in Del Negro and Schorfheide (2009), if one interprets the richer shock structure preferred by our method as a structural means of controlling for misspecification, the generalized shocks can simultaneously improve the accuracy of policy experiments and improve the fit of the model.

2.7 Appendix

2.7.1 Conventional Metropolis-Hastings Samplers

Markov Chain Monte Carlo (MCMC) methods in general provide samples from some probability distribution of interest by constructing a Markov chain whose stationary distribution is this distribution of interest. A Markov chain with the sequence of states $\varsigma_1, \varsigma_2, \ldots$ is specified in terms of the distribution for the initial state ς_1 and the transition kernel $K(\cdot)$ that provides the conditional distribution of a state ς_{i+1} given the current state ς_i. That is, the probability that ς_{i+1} is in some set $\mathcal{A} \subseteq \mathbb{R}^d$ given that the current state of the chain is ς_i is given by

$$K(\varsigma, \mathcal{A}) = P(\varsigma_{i+1} \in \mathcal{A} | \varsigma_i = \varsigma) \tag{2.20}$$

A distribution π is invariant for some Markov chain if the transition kernel of the chain satisfies

$$\int K(\varsigma, \mathcal{A}) \pi(\varsigma) d\varsigma = \int_{\mathcal{A}} \pi(\varsigma) d\varsigma \tag{2.21}$$

for all subsets $\mathcal{A}$ of the state space. The task in MCMC is to construct a kernel such that the distribution of interest π is invariant with respect to the Markov chain defined by $K()$. The expression in (2.21), however, is not practically useful for the construction of an appropriate kernel, as verifying (2.21) would involve integration over the unknown distribution π being sought.

One widely used approach to overcome this hurdle are Metropolis-Hastings samplers:[33] accept-reject samplers for which proposals for a new state of the chain are drawn from some distribution γ to be chosen by the researcher and then accepted with an appropriately derived probability α. Here, the stronger condition of reversibility or detailed balance is imposed, which guarantees that π is invariant for the Markov chain. This condition holds if a sequence of two

[33] Laid out in Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) and generalized in Hastings (1970).

states (ς, ς') has the same distribution as the reversed subchain (ς', ς) whenever $\varsigma, \varsigma' \sim \pi$. I.e., if

$$\int_{\mathcal{A}} \pi(\varsigma) K(\varsigma, \mathcal{B}) \, d\varsigma = \int_{\mathcal{B}} \pi(\varsigma') K(\varsigma', \mathcal{A}) \, d\varsigma' \tag{2.22}$$

for all subsets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^d$. Condition (2.22) is more easily verified and can thus provide a starting point for the construction of a sampler.

A general Metropolis-Hastings algorithm can be written as follows: Let again ς denote a state of the Markov chain, in the case of Bayesian inference in the context of model estimation, the state is just the vector of model parameters and the distribution of interest is the posterior distribution

$$\pi(\varsigma) \propto \mathcal{L}(\varsigma)\rho(\varsigma) \tag{2.23}$$

where ς denotes the vector of model parameters, $\mathcal{L}$ is the likelihood of the data given the model and its parameters and ρ is the prior over the model parameters. To obtain N samples from the posterior distribution, the following algorithm is run

Metropolis-Hastings

1. Set the (arbitrary) initial state ς_0 of the Markov chain
2. For $i = 1$ to N
   (a) Set $\varsigma = \varsigma_{i-1}$
   (b) Propose a new state from some proposal distribution $\gamma(\varsigma'|\varsigma)$
   (c) Accept draw with probability

   $$\alpha(\varsigma, \varsigma') = \min(1, \chi)$$

   with

   $$\chi = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \times \underbrace{\frac{\gamma(\varsigma|\varsigma')}{\gamma(\varsigma'|\varsigma)}}_{\text{Proposal Ratio}}$$

   (d) If the draw is accepted set $\varsigma_i = \varsigma'$. If the draw is rejected set $\varsigma_i = \varsigma$

This algorithm defines a transition kernel such that the Markov chain has the desired invariant distribution. The sequence of states of the chain is then a sample from this distribution of interest. The acceptance probability α corrects for differences between the proposal distribution γ and the distribution of interest.[34]

The kernel in the above is given by

$$K(\varsigma, \mathcal{B}) = \underbrace{\int_{\mathcal{B}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma'}_{\text{Probability of moving to set } \mathcal{B}} + \underbrace{\left[1 - \int_{\mathcal{B}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma'\right] \mathbb{1}_{\varsigma}}_{\text{Probability of rejecting the move and } \varsigma \in \mathcal{B}} \tag{2.24}$$

where $\mathbb{1}_{\varsigma} = 1$ if $\varsigma \in \mathcal{B}$ and zero otherwise, giving the probability of moving to some subset $\mathcal{B}$ of the parameter space conditional on the chain currently being at ς. The crux when constructing the kernel is to define the appropriate acceptance probability α and the proposal distribution γ so as to satisfy the detailed balance condition and thereby guarantee the convergence of the Markov chain to the desired probability distribution. Indeed, plugging in the formulation of the

[34] Note that in the case of a standard random walk Metropolis-Hastings sampler with symmetric proposals, i.e. a Metropolis sampler, the proposal ratio reduces to one.

kernel from (2.24) into (2.22) gives an expression from which, given the proposal distribution γ the appropriate acceptance probability α can be readily derived using Peskun's (1973) recipe.

2.7.2 Detailed Derivation of Inflated Proposal Mapping

To choose an appropriate mapping $g_{pp'}$, it is useful to break the mapping into two parts according to the desired parameters P_p and the auxiliary parameters u. The mapping $g_{pp'}$ is given by

$$(P^{p'}, u') = g_{pp'}(P^p, u) = (g_{1pp'}(P^p, u), g_{2pp'}(P^p, u)) \tag{2.25}$$

and its inverse

$$(P^p, u) = g_{pp'}^{-1}(P^{p'}, u') = g_{p'p}(P^{p'}, u') = (g_{1p'p}(P^{p'}, u'), g_{2p'p}(P^{p'}, u') \tag{2.26}$$

Start with $g_{1pp'}$. Suppose now that the current state of the Markov chain is at $\varsigma = (p, P^p)$. Now with probability $\gamma_p(p'|p)$, a move to the model with order p' is proposed. Conditional on this proposal, we draw u from some proposal distribution $\gamma_{pp'}(u)$. Then, we introduce a deterministic mapping $g_{1pp'}$ that maps the current state and the auxiliary proposal u to the proposed new state such that $(p', P^{p'}) = (p', g_{1pp'}(P^p, u))$. Note that u is not part of the state of the chain.

Additionally, we have to find $g_{2pp'}$. In order to be able to easily verify adherence to detailed balance for a move from a state (p, P^p) to $(p', P^{p'}) = (p', g_{1pp'}(P^p, u))$ the vectors of Markov chain states and the random auxiliary proposal variables (P^p, u) and $(P^{p'}, u')$ must be of equal dimension and requiring $g_{pp'}$ to be a differentiable bijection lets us use a simple change-of-variables in the detailed balance equation. I.e., the kernel of the chain is now defined in terms of the auxiliary variable u together with the model indicator and the parameter vectors.

Armed with this structure it is now straightforward to derive the appropriate acceptance probability. The detailed balance condition holds if[35]

$$\int_{\mathcal{A}_p} \pi\left(p|y\right) \pi\left(P^p|p,y\right) Q\left(\varsigma, \mathcal{B}_{p'}\right) dP^p = \int_{\mathcal{B}_{p'}} \pi\left(p'|y\right) \pi\left(P^{p'}|p',y\right) Q\left(\varsigma', \mathcal{A}_p\right) dP^{p'} \tag{2.27}$$

for all subsets $\mathcal{A}_p$ and $\mathcal{B}_{p'}$ of the parameter spaces associated with autoregressive polynomials of order p and p' respectively. The posterior distribution $\pi(\varsigma|y)$ is factorized as $\pi(\varsigma|y) = \pi(p|y)\pi(P^p|p,y)$ and

$$\begin{aligned} Q\left(\varsigma, \mathcal{B}_{p'}\right) &= \int_{\mathcal{B}_{p'}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma,\varsigma')d\varsigma' \\ &= \gamma_p(p'|p) \int \mathbb{1}(g_{1pp'}(P^p,u) \in \mathcal{B}_{p'})\alpha_{pp'}(P^p, g_{1pp'}(P^p,u)\gamma_{pp'}(P^p,u)du \end{aligned}$$

The left hand side of (2.27) is then

$$\begin{aligned} \int_{\mathcal{A}_p} \pi\left(\varsigma|y\right) Q\left(\varsigma, \mathcal{B}_{p'}\right) dP^p = \int\int \mathbb{1}(P^p \in \mathcal{A}_p, g_{1pp'}(P^p,u) \in \mathcal{B}_{p'})\pi\left(p|y\right)\pi\left(P^p|p,y\right) \times & \qquad (2.28) \\ \gamma_p(p'|p)\alpha_{pp'}(P^p, g_{1pp'}(P^p,u)\gamma_{pp'}(P^p,u)dP^p du & \qquad (2.29) \end{aligned}$$

and the right hand side reads

$$\begin{aligned} \int_{\mathcal{B}_{p'}} \pi\left(\varsigma'|y\right) Q\left(\varsigma', \mathcal{A}_p\right) dP^{p'} = \int\int \mathbb{1}(P^{p'} \in \mathcal{B}_{p'}, g_{1p'p}(P^{p'},u') \in \mathcal{A}_p)\pi\left(p'|y\right)\pi\left(P^{p'}|p',y\right) \times & \qquad (2.30) \\ \gamma_p(p|p')\alpha_{p'p}(P^{p'}, g_{1p'p}(P^{p'},u'))\gamma_{p'p}(P^{p'},u')dP^{p'} du' & \qquad (2.31) \end{aligned}$$

where $\gamma(\varsigma'|\varsigma)$ is again factorized as $\gamma_p(p^{|}p)\gamma_{pp'}(P^p,u)$. The fact that $g_{pp'}$ is a differentiable bijection together with the dimension matching conditions enables a change of variable in

[35]See also Waagepetersen and Sorensen (2001).

(2.30) leading to

$$\begin{aligned}&\int\int 1(g_{1pp'}(P^p,u)\in\mathcal{B}_{p'},P^p\in\mathcal{A}_p)\pi\left(p'|y\right)\pi\left(g_{1pp'}(P^p,u)|p',y\right)\gamma_p(p|p')\\ &\times\alpha_{p'p}(g_{1pp'}(P^p,u),P^p)\gamma_{p'p}(g_{1pp'}(P^p,u),g_{2pp'}(P^p,u))|g'_{pp'}(P^p,u)|dP^p du\end{aligned}\tag{2.32}$$

where $dP^{p'}du' = |g'_{pp'}(P^p,u)|dP^p du$ and $|g'_{pp'}(P^p,u)|$ is the determinant of the Jacobian of $g_{pp'}$.

By inspection of (2.28) and (2.32), the reversibility condition (2.27) is satisfied if

$$\begin{aligned}&\pi\left(p|y\right)\pi\left(P^p|p,y\right)\gamma_p(p'|p)\alpha_{pp'}(P^p,g_{1pp'}(P^p,u))\gamma_{pp'}(P^p,u) =\\ &\qquad\pi\left(p'|y\right)\pi\left(g_{1pp'}(P^p,u)|p',y\right)\gamma_p(p|p')\alpha_{p'p}(g_{1pp'}(P^p,u),P^p)\times\\ &\qquad\gamma_{p'p}(g_{1pp'}(P^p,u),g_{2pp'}(P^p,u))|g'_{pp'}(P^p,u)|\end{aligned}\tag{2.33}$$

Choosing the acceptance probability as large as possible, we have

$$\alpha_{pp'} = \min\left(1,\chi_{pp'}(\varsigma,\varsigma')\right)\tag{2.34}$$

with

$$\chi_{pp'}\left(\varsigma,\varsigma'\right) = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}}\underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}}\underbrace{\frac{\gamma_p(p|p')\gamma_{p'p}(g_{pp'}(P^p,u))}{\gamma_p(p'|p)\gamma_{pp'}(P^p,u)}|g'_{pp'}\left(P^p,u\right)|}_{\text{Proposal Ratio}}\tag{2.35}$$

With our mapping $g_{pp'}$, in (2.4), $|g'_{pp'}(P^p,u)|$ is equal to one and (2.35) reduces to (2.9).[36]

[36] The posterior π is here written factorized as the product of likelihood and prior $\mathcal{L}(\varsigma)\rho(\varsigma)$ for correspondence with the general formulation of the detailed balance condition (2.22).

2.7.3 Imposing Stationarity and Invertibility on ARMA(p,q) Sampling

To constrain sampling to these invertible and stationary regions of the parameters spaces of each model, we follow Barndorff-Nielsen and Schou (1973), Monahan (1984) and Jones (1987) and reparametrize the AR (and MA) polynomial in terms of its (inverse) partial autocorrelations (PACs). If the (inverse) partial autocorrelations are between -1 and 1 the process is (invertible) stationary.

First, we generalize the AR(p) model to an ARMA(p,q) as follows

(2.36)

$$y_t = P_1^{p,q} y_{t-1} + P_2^{p,q} y_{t-2} + \ldots + P_p^{p,q} y_{t-p} + \epsilon_t + Q_1^{p,q}\epsilon_{t-1} + \ldots + Q_q^{p,q}\epsilon_{t-q}, \;\; \epsilon_t \sim \mathcal{N}\left(0, \sigma^2\right)$$

In order to recover the coefficients of the AR polynomials, the following algorithm is run

Recovering AR Coefficients from PACs

1. Introduce $p^k = \left(p_1^{(k)}, \ldots, p_k^{(k)}\right), k = 1, \ldots, p$
2. Draw $r = r_1, \ldots, r_p,$ for $r_i \in (0, 1)$ partial autocorrelations
3. Set $p_1^{(1)} = r_1$
4. Run the recursion

$$p_i^{(k)} = p_i^{(k-1)} - r_k p_{k-i}^{(k-1)}, i = 1, \text{ for } \ldots, k-1$$

   with $p_k^{(k)} = r_k$ for $k = 2, \ldots, p$
5. Set $P^p = p^{(p)}$

The MA coefficients are recovered analogously, where the inverse partial autocorrelations substitute for the partial autocorrelations, r_i, in the foregoing. Ultimately, instead of proposing AR(MA) parameters directly, (inverse) partial autocorrelations are proposed in their place from which the parameters are then recovered. This will obviously necessitate the formulation of priors over (inverse) partial autocorrelations instead of parameters.

Chapter 3

Solving and Estimating Linearized DSGE Models with VARMA Shock Processes and Filtered Data

Solving and Estimating Linearized DSGE Models with VARMA Shock Processes and Filtered Data *

Alexander Meyer-Gohde[†§]   Daniel Neuhoff [‡]

Abstract

We derive recursive solutions to linearized DSGE models with VARMA exogenous driving forces of arbitrary order without inflating the state vector. Representing the solution in the frequency domain, we calculate the likelihood of a sequence of observations from the model, as well as its nonrecursively filtered (e.g., Hodrick-Prescott or Baxter-King) variant straightforwardly.

JEL classification: C51, C61, C63, E17

Keywords: DSGE models; ARMA; VAR; likelihood function;

*We are grateful to Michael Burda for useful comments. This research was supported by the DFG through the SFB 649 "Economic Risk." Any and all errors are entirely our own.

†Universität Hamburg, Professur für Volkswirtschaftslehre insb. Wachstum und Konjunktur, Von-Melle-Park 5, 20146 Hamburg, Germany; Tel.: +49-40-42838 3996; E-Mail: alexander.meyer-gohde@wiso.uni-hamburg.de

§Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5720; E-Mail: alexander.meyer-gohde@wiwi.hu-berlin.de

‡Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5667; E-Mail:daniel.neuhoff@gmail.com

3.1 Introduction

Many DSGE models relax the standard independent AR(1) assumption on exogenous processes.[1] Sims (2001, p. 4) notes that "common practice" is to model serial correlation by incorporating the exogenous processes into the endogenous state vector. This increases the burden of solving the matrix quadratic problem and evaluating the likelihood function, especially with complicated patterns of serial correlation. The literature for solving DSGE models has not yet provided an implementable method for solving models with VARMA driving processes of arbitrary order without resorting to inflating the endogenous state vector.[2] We close this gap by leaving the exogenous processes out of the endogenous state vector and solve for the unknown coefficients associated with the inhomogenous solution by solving a series of linear (generalized Sylvester) equations.

Schmitt-Grohé and Uribe (2010) note that when the model is to be treated consistently with the data with a nonrecursive filter, such as the Hodrick and Prescott (1997) (HP) filter or a band-pass (BP) filter (see, e.g., Baxter and King (1999)), the Kalman filter cannot be used to evaluate the likelihood. Schmitt-Grohé and Uribe (2010), however, proceed with their calculations in the time domain, precluding the application of such filters. We follow their approach, treating the sample as a single draw from a multivariate normal distribution, but calculate the autocovariances spectrally to enable us to explicitly apply the HP or BP filter to the model while evaluating the likelihood.

---

[1] E.g., ARMA(1,1) is used for Smets and Wouters's (2007) markup shocks and Croce's (2014) long-run growth, AR(2) for Del Negro and Schorfheide's (2009) government expenditures, and VAR(6) in (Cúrdia and Reis 2010).

[2] E.g., Sims (2001) and Anderson (2010), provide formulas applicable to arbitrary exogenous processes, yet these are left as discounted expected sums. In essence, we evaluate these sums here for VARMA(p,q) processes.

3.2 Model

We consider linear(ized) DSGE models expressed as

$$\underset{n_x\times 1}{0} = E_t\left[A\underset{n_x\times 1}{X_{t+1}} + BX_t + CX_{t-1} + D\underset{n_z\times 1}{Z_t}\right] \tag{3.1}$$

where the vector X_t collects the endogenous variables and the vector Z_t the exogenous variables.

We admit arbitrary stationary VARMA(p,q) processes for Z_t

$$Z_t = P_1 Z_{t-1} + P_2 Z_{t-2} \ldots + P_p Z_{t-p} + \epsilon_t + Q_1\epsilon_{t-1} \ldots + Q_q\epsilon_{t-q}, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma) \tag{3.2}$$

where stationarity follows from the assumption on :$P(\lambda) \doteq I_{n_z}\lambda^p - P_1\lambda^{p-1} + P_2\lambda^{p-2} \ldots + P_p$:

Assumption 3.2.1. $\exists\, n_z$, *counting multiplicities,* $\lambda \in \mathbb{C} : \det P(\lambda) = 0, |\lambda| < 1.$

3.3 Solution

Given (3.2) and (3.1), the state variables of the model, where $\tilde{p} = \max(p, 1)$,[3] are

$$\left\{X_{t-1}, Z_t, Z_{t-1}, \ldots, Z_{t-(\tilde{p}-1)}, \epsilon_t, \epsilon_{t-1}, \ldots, \epsilon_{t-(q-1)}\right\} \tag{3.3}$$

The solution with unknown coefficients $\{\Lambda, \Phi_0, \Phi_1, \ldots, \Phi_{\tilde{p}-1}, \Theta_0, \Theta_1, \ldots, \Theta_{q-1}\}$ is

(3.4)

$$X_t = \Lambda X_{t-1} + \Phi_0 Z_t + \Phi_1 Z_{t-1} \ldots + \Phi_{\tilde{p}-1} Z_{t-(\tilde{p}-1)} + \Theta_0\epsilon_t + \Theta_1\epsilon_{t-1} \ldots + \Theta_{q-1}\epsilon_{t-(q-1)}$$

[3] This follows from (3.2) expressed in first order form for appropriate PP and QQ matrices

$$\begin{bmatrix} Z_t' & Z_{t-1}' & \ldots & Z_{t-(\tilde{p}-1)}' & \epsilon_t' & \epsilon_{t-1}' & \ldots & \epsilon_{t-(q-1)} \end{bmatrix}' = PP\begin{bmatrix} Z_{t-1}' & Z_{t-2}' & \ldots & Z_{t-p}' & \epsilon_{t-1}' & \epsilon_{t-1}' & \ldots & \epsilon_{t-q} \end{bmatrix}' + QQ\epsilon_t$$

The left hand side is the exogenous state vector. The case $p = 0$ is permitted through $\tilde{p}$.

Solving for Λ is standard in the literature,[4] we simply assume a unique stable solution

Assumption 3.3.1. *$\exists$ a unique $\Lambda \in \mathbb{R}^{n_x \times n_x} : A\Lambda^2 + B\Lambda + C = 0, |eig(\Lambda)| < 1.$*

Our assumptions give the model (3.1) a unique stable solution, summarized by:

Proposition 3.3.2. *Let assumptions 3.2.1 and 3.3.1 hold. There exists a unique stable solution (3.4) to (3.1). The solvent of assumption 3.3.1 is Λ, $\{\Theta_0, \Theta_1, \ldots, \Theta_{q-1}\}$ solve*

$$\underset{n_x \times n_z}{0} = A\left(\Lambda\Theta_i + \Phi_0 Q_{i+1} + \Theta_{i+1}\right) + B\Theta_i, \text{ for } i = 0, \ldots, q-1 \tag{3.5}$$

where $\Theta_i = \underset{n_x \times n_z}{0} \; \forall i > q-1$, and $\{\Phi_0, \Phi_1, \ldots, \Phi_{p-1}\}$ solve

$$\begin{aligned} \underset{n_x \times n_z}{0} &= A\left(\Lambda\Phi_0 + \Phi_0 P_1 + \Phi_1\right) + B\Phi_0 + D, \text{ and} \\ \underset{n_x \times n_z}{0} &= A\left(\Lambda\Phi_i + \Phi_0 P_{i+1} + \Phi_{i+1}\right) + B\Phi_i, \text{ for } i = 0, \ldots, p-1 \end{aligned} \tag{3.6}$$

for $p > 1$, where $\Phi_i = \underset{n_x \times n_z}{0} \; \forall i > p-1$, and Φ_0 solves

$$\underset{n_x \times n_z}{0} = A\left(\Lambda\Phi_0 + \Phi_0 P_1\right) + B\Phi_0 + D \tag{3.7}$$

otherwise, where $P_1 = \underset{n_z \times n_z}{0}$ if $p = 0$.

Proof. See the appendix. □

To solve $\{\Phi_i\}_{i=0}^{\tilde{p}-1}$, we rewrite (3.6), or (3.7) if $p \leq 1$, as

$$\Phi_{p-i} = \begin{cases} \sum_{j=1}^{i} \left(-\left(B + A\Lambda\right)^{-1} A\right)^j \Phi_0 P_{p-i+j} & \text{for } i = 1, 2, \ldots p-1 \\ \sum_{j=1}^{i} \left(-\left(B + A\Lambda\right)^{-1} A\right)^j \Phi_0 P_{p-i+j} - \left(B + A\Lambda\right)^{-1} D & \text{for } i = p \end{cases} \tag{3.8}$$

where the invertibility of $B + A\Lambda$ follows from assumption 3.3.1, see Lan and Meyer-Gohde

[4] E.g., Klein (2000) or Anderson (2010) for higher endogenous leads and lags.

(2012).

For $i = p$, (3.8) is $\Phi_0 + \sum_{j=1}^{p} \left(-\left(B + A\Lambda\right)^{-1} A\right)^j \Phi_0 \left(-P_j\right) = -\left(B + A\Lambda\right)^{-1} D$, a linear equation in Φ_0 (a p'th generalized Sylvester equation) solved in the appendix.

Given Φ_0, $\{\Phi_i\}_{i=1}^{p-1}$ follow from (3.6) starting with Φ_{p-1} and working backwards to Φ_1. Given Φ_0, $\{\Theta_i\}_{i=0}^{q-1}$ follow from (3.5) starting with Θ_{q-1} and working backwards to Θ_0.

The recursive solution (3.4) possess a MA(∞) representation, summarized in:

Proposition 3.3.3. *Let assumptions 3.3.1 and 3.2.1 hold. The unique, stable solution (3.4) to (3.1) for X_t in proposition 3.3.2 has a unique infinite moving average representation given by*

$$X_t = \Big(\underset{n_x \times n_x}{I} - \Lambda L\Big)^{-1} \left[\Phi\left(L\right) P\left(L\right)^{-1} Q\left(L\right) + \Theta\left(L\right)\right] \epsilon_t \tag{3.9}$$

where L is the lag operator (e.g., Sargent (1987)) and

$$P\left(L\right) \doteq I_{n_z} - P_1 L - P_2 L^2 \ldots - P_p L^p, \quad Q\left(L\right) \doteq I + Q_1 L \ldots + Q_q L^q \tag{3.10}$$

$$\Phi\left(L\right) \doteq \Phi_0 + \Phi_1 L \ldots + \Phi_{\tilde{p}-1} L^{\tilde{p}-1}, \quad \Theta\left(L\right) \doteq \Theta_0 + \Theta_1 L \ldots + \Theta_{q-1} L^{q-1} \tag{3.11}$$

Proof. See the appendix. □

3.4 Likelihood

We apply filters to the model while evaluating the likelihood by calculating the sequence of autocovariances spectrally. We provide two examples: the HP filter and a BP filter.

A linear combination of elements of X_t, e.g., observables, possess the MA(∞) representation

$$Y_t = \underset{n_y \times n_x}{\Upsilon^X}, \;\; X_t = \Upsilon^X \Big(\underset{n_x \times n_x}{I} - \Lambda L\Big)^{-1} \left[\Phi\left(L\right) P\left(L\right)^{-1} Q\left(L\right) + \Theta\left(L\right)\right] \epsilon_t \tag{3.12}$$

E.g., Sargent (1987) or Uhlig (1999) show the autocovariances of Y_t, $\Gamma_n \doteq E\left[Y_t Y_{t-n}'\right]$, are

$$\Gamma_n = \int_{-\pi}^{\pi} G(\omega) e^{i\omega n} d\omega \tag{3.13}$$

the inverse Fourier transformation of the spectral density of Y_t, $G(\omega)$ given by

(3.14)

$$G(\omega) \doteq H(-\omega)\Sigma H(\omega)', \quad H(\omega) = \Upsilon^X \left(\underset{n_x \times n_x}{I} - \Lambda e^{i\omega}\right)^{-1} \left[\Phi\left(e^{i\omega}\right) P\left(e^{i\omega}\right)^{-1} Q\left(e^{i\omega}\right) + \Theta\left(e^{i\omega}\right)\right]$$

Applying the HP filter to the model, we can use closed form representation of the HP filter in the frequency domain, see King and Rebelo (1993), given as

$$HP(\lambda, \omega) = 4\lambda\left(1 - \cos(\omega)\right)^2 / \left[1 + 4\lambda\left(1 - \cos(\omega)\right)^2\right] \tag{3.15}$$

where λ is the smoothing parameter and ω a frequency or the band-pass filter

$$BP(\underline{\omega}, \overline{\omega}, \omega) = 1 \text{ if } |\omega| \in [\underline{\omega}, \overline{\omega}] \text{ and } 0 \text{ otherwise} \tag{3.16}$$

where $\underline{\omega}$ and $\overline{\omega}$ define the lower and upper bounds on frequencies not removed by the filter. In either case, the autocovariances of the filtered observables are

$$\Gamma_n = \int_{-\pi}^{\pi} F(\omega)^2 G(\omega) e^{i\omega n} d\omega \tag{3.17}$$

where $F(\omega) = HP(\lambda, \omega)$ for the HP and $F(\omega) = BP(\underline{\omega}, \overline{\omega}, \omega)$ for the band-pass filter.

T observations of Y_t, $Y = [Y_1' Y_2' \ldots Y_T']'$, are then normal with block Toeplitz covariance matrix

$$\Psi = \begin{bmatrix} \Gamma_0 & \Gamma_1' & \dots & \Gamma_{T-1}' \\ \Gamma_1 & \Gamma_0 & \dots & \Gamma_{T-2}' \\ \vdots & & \ddots & \vdots \\ \Gamma_{T-1} & & \dots & \Gamma_0 \end{bmatrix} \tag{3.18}$$

with autocovariances, Γ_n, from (3.13) or (3.17); the log-likelihood of parameters ς given the data is

$$\mathcal{L}(\varsigma|Y) = -0.5pTln\left(2\pi\right) - 0.5ln\left(det\left(\Psi(\vartheta)\right)\right) - 0.5Y'\Psi(\vartheta)^{-1}Y \tag{3.19}$$

$ln\left(det\left(\Psi(\vartheta)\right)\right)$ and $Y'\Psi(\vartheta)^{-1}Y$ can be calculated with (3.18) following, e.g., Meyer-Gohde (2010).

3.5 Conclusion

We have provided a solution for linear DSGE models with general VARMA(p,q) exogenous processes without expanding the endogenous state vector. We calculated the autocovariances spectrally, enabling the application of nonrecursive filters while calculating the likelihood.

3.6 Appendix

3.6.1 Proof of Proposition 3.3.2

Insert (3.4) for X_t once and for X_{t+1} twice in (3.1), substitute (3.2) lagged forward for the Z_{t+1} that arises when X_{t+1} is replaced with (3.4). The coefficients on (3.3) are zero, as the solution (3.4) must hold in all states, delivering the equations in the proposition. Stability follows from assumptions 3.2.1 and 3.3.1.

3.6.2 Generalized Sylvester Equations

The equation for Φ_0 takes the form

(3.20)
$$x + \beta x \gamma_1 + \beta^2 x \gamma_2 \ldots + \beta^p x \gamma_J = \delta$$

where $x \doteq \Phi_0$, $\beta \doteq -(B + A\Lambda)^{-1} A$, $\gamma_j \doteq -P_j$, for $j = 1, 2, \ldots, p$, and $\delta \doteq -(B + A\Lambda)^{-1} D$.

Proposition 3.6.1. *The generalized Sylvester equation* $\underset{n_a \times n_b}{x} + \beta x \gamma_1 + \beta^2 x \gamma_2 \ldots + \beta^J x \gamma_J = \delta$ *can be solved recursively for* $\tilde{x} \doteq Q^\dagger x$ *starting with the last (n_a'th) row of* $\tilde{x}$ *as follows*

(3.21)
$$\tilde{x}_{i,\bullet} = \left[\delta_{i,\bullet} - \left(\sum_{k=1}^{n_a - i} \sum_{j=0}^{J} \{U^j\}_{i, n_a + k} \tilde{x}_{n_a + k, \bullet} \gamma_j \right) \right] \left(\sum_{j=0}^{J} \gamma_j U_{i,i}^j \right)^{-1}, \quad \text{for } i = n_a, n_a - 1, \ldots, 1$$

where $QUQ^\dagger = \beta$ *with* U *upper diagonal and* Q *unitary is the complex Schur decomposition of* β*, conjugate transposition is denoted by* †*, and* $_{c,d}$ *references the* c*'th row and* d*'th column.*

Proof. With the Schur decomposition $QUQ^\dagger = \beta$ and noting that $Q^\dagger = Q^{-1}$, (3.20) is $x + QUQ^\dagger x \gamma_1 + QU^2 Q^\dagger x \gamma_2 \ldots + QU^J Q^\dagger x \gamma_J = \delta$, multiplying through with $Q^\dagger$ and using the definition $\tilde{x} \doteq Q^\dagger x$ gives $\tilde{x} + U \tilde{x} \gamma_1 + U^2 \tilde{x} \gamma_2 \ldots + U^J \tilde{x} \gamma_J = Q^\dagger \delta$. As U is upper diagonal, so is any power of U; thus given all rows of the matrix $\tilde{x}$ after some i, the i'th row of $\tilde{x}$, $\tilde{x}_{i,\bullet}$ solves $\sum_{j=0}^{J} U_{i,i}^j \tilde{x}_{i,\bullet} \gamma_j = \delta_{i,\bullet} - \left(\sum_{k=1}^{n_a - i} \sum_{j=0}^{J} \{U^j\}_{i, n_a + k} \tilde{x}_{n_a + k, \bullet} \gamma_j \right)$ and noting that $U_{i,i}$ is a scalar and right-multiplying with the inverse of $\sum_{j=0}^{J} \gamma_j U_{i,i}^{j}$ gives (3.21). $\square$

3.6.3 Proof of Proposition 3.3.3

Invertibility of $\left(\underset{n_x \times n_x}{I} - \Lambda L \right)$ follows from proposition 3.3.2 and that of $P(L)$ from lemma 3.2.1. Uniqueness follows from assumption 3.3.1 and from proposition 3.3.2.

Chapter 4

Dynamics of Real Per Capita GDP

Dynamics of Real Per Capita GDP *

Daniel Neuhoff †

Abstract

This study investigates the dynamics of quarterly real GDP per capita growth rates across four countries, the US, UK, Canada and France. I obtain estimates for ARIMA(p,q) processes for first differences of log quarterly real GDP per capita using Reversible Jump Markov Chain Monte Carlo, allowing me to account for model uncertainty when comparing the implied impulse responses across countries.

The estimated impulse response functions differ in shape. The persistence estimates for the US, France, Canada and Italy are clustered together, while the UK and Japan are clear outliers. Substantial posterior uncertainty remains regarding the persistence estimates and the appropriate ARMA models. The countries are ranked according to estimated persistenc. This ranking is robust with respect to the detrending device employed. The results for the UK depend on the time period considered. An analysis of the components of GDP for the US suggests that the dynamics are mainly driven by consumption.

JEL classification: C51, C52, E32

Keywords: ARMA; Real GDP per capita; Growth Rates; Persistence; Reversible Jump Markov Chain Monte Carlo

---

*I am grateful to Michael Burda and Alexander Meyer-Gohde for useful comments. Furthermore, I am indebted to Tobias König for excellent research assistance. This research was supported by the DFG through the SFB 649 "Economic Risk." Any and all errors are entirely my own.

†Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5667; E-Mail:daniel.neuhoff@gmail.com

4.1 Introduction

The dynamic behavior of GDP has attracted longstanding interest among economists. This study aims to add to the existing literature by investigating the dynamics of quarterly real GDP per capita across six countries, the US, UK, Canada, Italy, Japan, and France. In contrast to earlier studies on the dynamics of growth rates, such as Campbell and Mankiw (1987) who obtain maximum likelihood point estimates for ARIMA(p,1,q) models of quarterly real GNP in the US, this investigation employs Reversible Jump Markov Chain Monte Carlo (henceforth RJMCMC).

This Bayesian approach enables the sampling from posteriors across models where the associated parameter spaces vary in dimensionality from model to model. The posterior will then not only incorporate posterior uncertainty about parameter values, as is the case for fixed-dimension Bayesian methods like Random-Walk MCMC, but also reflects posterior uncertainty about the models themselves while at the same time providing a method to efficiently traverse the model space.

I analyze the posterior distributions of the impulse responses as well as the measure of persistence based on cumulative impulse responses also utilized in Campbell and Mankiw (1989). The results are compared to maximum likelihood estimates with model choice according to three information criteria: Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC), and the Bayesian Information Criterion (BIC). The results from maximum likelihood estimates mostly coincide with the means and modes of the posterior impulse responses when the model is chosen using the BIC. In contrast, both AIC and AICC choose less parsimonious models exhibiting much higher persistence and often oscillatory behavior of the impulse responses, where the latter is rare for estimates using either RJMCMC or the BIC.

The comparison of impulse responses across countries also reveals significant variation: for the UK, RJMCMC assigns an extensive amount of posterior mass to the pure random walk model for the level of GDP. The impulse response function of a shock to the growth rate of GDP in the UK therefore exhibits very little persistence. In contrast, the posterior impulse responses for the growth rates of Canada and the US exhibit more persistence with the median responses dying out after 5 to 6 quarters. France and Italy show somewhat higher persistence, while Japan is consistently ranked as the economy with the most persistent response to a shock. It is shown, however, that the results for the UK do not carry over when the time series for UK GDP is split into subsamples.

For the posterior distributions of the persistence measure, two-sample Kolmogorov-Smirnov tests are carried out. In all cases, the null hypothesis that any combination of posteriors is generated by the same distribution is rejected at the 1% level.

In order to gain some insight into which component of GDP drives the results, the method is applied to the major aggregates of US GDP– private and government consumption, as well as fixed capital formation, exports, and imports. The results suggest, that the shape of the impulse response for the GDP series is mainly driven by private and government consumption.

Since it is well known that the detrending method chosen has significant impact on empirical results, see e.g. Canova (1998), the results from the difference stationary perspective are compared with the results obtained using linear or Ordinary Least Squares (OLS) and Hodrick-Prescott (HP)[1] detrending. The results for the HP filtered series seem to be dominated by filtering artifacts, while the results for linear detrending are in line with the ones from the difference stationary perspective.

In general, distinguishing a trend-stationary process with a large autoregressive root from a unit root process seems unfeasible with available data as emphasized by Christiano and Eichenbaum (1990), among others, who state that "[to] us the possibility of providing a compelling case that real GNP is either trend or difference stationary seems extremely small". Furthermore, in their seminal contribution Kwiatkowski, Phillips, Schmidt, and Shin (1992) find that for real GNP per capita they "cannot reject either the unit root hypothesis or the trend stationary hypothesis".

The results suggest that economic models that put strong constraints on the dynamic

[1] See Hodrick and Prescott (1997).

response of GDP growth rates to reduced form shocks, may only be appropriate in certain instances. Furthermore, the dynamics may change significantly over time as suggested by the results for subsamples of UK GDP. For the US, the dynamics appear to be driven mainly by government consumption, private consumption, and to a lesser extent, investment.

The rest of this chapter is structured as follows: after a review of some of the relevant literature, I discuss the relationship between point estimates and posterior distributions, setting up a brief discussion on the estimation of ARMA models with RJMCMC and the frequentist approach employed here. After a discussion of the data and the sampler settings, a measure of persistence is introduced, in order to then present the results for GDP growth rates and the robustness check. Following the persistence results, I discuss the results from the GDP components and subsamples from the UK and end with a conclusion.

4.2 Literature

The study of the dynamic properties of output measures has inspired longstanding substantial interest among economists. The strand of literature bearing the closest resemblance to the investigation presented here was initiated by Campbell and Mankiw (1987) who analyze the persistence of US GNP from a difference stationary perspective after Nelson and Plosser (1982) had challenged the hitherto prevailing view among economists that aggregate time series were trend stationary. Other studies concerned with trends in and persistence of output and other economic variables include Clark (1989), Stock and Watson (1988), and Watson (1986). While the researchers disagree on the long-run effect of an innovation, there is cautious consensus that significant persistence is present in economic time series.

Campbell and Mankiw (1989) provide an international perspective on persistence in a difference stationary world, confirming the finding of meaningful levels of persistence for the G7 economies. Among others, Cochrane (1988) challenges the view that GNP is clearly difference stationary. Using Bayesian techniques, DeJong and Whiteman (1991) find significant support for time trends in the posterior distributions for many of the series analyzed by Nelson and Plosser (1982). Perron (1993) finds that when allowing for a break in the trend, trend stationarity seems to be a good description of the behavior of the data. Perron's paper was, however, criticized for picking the break point in the trend a priori. Cheung (1994) carries out unit root tests allowing the structural break to be determined endogenously and rejects the null hypothesis of a unit root. He finds significant differences in the dynamic behavior of GDP across countries, which is consistent with the conclusions of Campbell and Mankiw (1989). Koop (1991) analyzes the time series properties of real per capita GDP for 121 countries using a Bayesian approach confirming the results from previous studies with respect to persistence. He finds mixed evidence regarding the trend and difference stationarity hypotheses.

While trend stationary and difference stationary models offer extremely differing implications, especially concerning long term forecasts, the question of which model is closest to the true nature of GDP is unlikely to be settled in the near future, as also argued in Christiano and Eichenbaum (1990). Hence, both perspectives will be considered in this chapter.

Another strand of literature concerned with breaking up the dichotomy between trend and difference stationarity uses fractionally integrated time series models to analyze output series. Studies in this vein include Diebold and Rudebusch (1989) and Cheung and Lai (1992). Both studies find substantial persistence in output, albeit not always unit roots, for the countries considered.

While univariate time series analyses of output appear to be simple or even naive, the resulting findings can be used, for example, to analyze the welfare implications of stabilization policies as well as discriminate between economic models as in Durlauf, Romer, and Sims (1989). Other authors such as Jones (1995), Ragacs and Zagler (2002), Fatas (2000b), and Fatas (2000a) use univariate results to test models of economic growth. Furthermore, multivariate econometric models have univariate representations, as pointed out already by Quenouille (1957), and DSGE models in turn possess VAR representations - of finite or infinite order - as shown by Ravenna (2007).

4.3 Point Estimates vs. Posterior Distributions

In the following, results from a Bayesian approach to time series estimation are compared to their frequentist counterparts. Apart from their philosophical differences with respect to conditioning, the two approaches also give different output: the frequentist approach yields *point estimates* of parameters together with confidence intervals around these estimates which are then compared to some *limiting distribution* of the estimator for inference, while the Bayesian approach delivers *posterior distributions* of the parameters on which inference is based.

Based on these distributions, point estimates for parameters and any function of the parameters can be derived by choosing a loss function. Loss functions are in essence penalties for "missing" the true parameter values. This is akin to minimizing the sum of squares of the deviations of the data from their model-implied value in a classical linear regression. Commonly used are the quadratic loss function yielding the mean of the posterior distribution as the estimator, and the absolute loss giving the median of the posterior as estimate. Throughout the following, complete posterior distributions will be compared with each other using the two-sample Kolmogorov-Smirnov-Test as well as the point estimates for the two commonly used loss functions together with credible sets.

4.4 Bayesian Estimation of ARMA Models Using RJMCMC

The estimation carried out here employs the Reversible Jump Markov Chain Monte Carlo (RJMCMC) methodology pioneered by Green (1995). RJMCMC generalizes the Metropolis-Hastings algorithm from Hastings (1970) to allow sampling from posterior distributions spanning different models and therefore parameter spaces of variable dimensionality. The method is applied here to obtain posterior distributions spanning the model and corresponding parameter spaces of stationary and invertible ARMA(p,q) models with $p, q \in [0; 10]$ of the form:

$$P(L)y_t = Q(L)\epsilon_t; \epsilon_t \sim N(0, \sigma_e^2) \tag{4.1}$$

with

$$P(L) = 1 - P_1 L - P_2 L^2 - ...P_p L^p \tag{4.2}$$

$$Q(L) = 1 + Q_1 L + Q_2 L^2 + ...Q_q L^q \tag{4.3}$$

denoting the autoregressive and moving average polynomials respectively and L denoting the lag operator. It is assumed throughout that the coefficients of $Q(L)$ satisfy the invertibility and those of $P(L)$ the stationarity conditions. In order to impose these conditions, the model is reparametrized in terms of its (inverse) partial autocorrelations for the (moving average) autoregressive polynomials as in e.g. Barndorff-Nielsen and Schou (1973), Monahan (1984), Jones (1987), and chapters 2 and 5 of this thesis. These assumptions as well as the notation will be used throughout this study.

The RJMCMC implementation employed here is identical to that in chapter 2 including the evaluation of the likelihood by means of the Kalman filter, apart from the fact that two proposal distributions are used, one for within-model moves where the model indicators remain constant and one for model moves where at least one indicator changes. An in-depth explanation of and references to other literature about the RJMCMC algorithm applied here can be found in chapter 2.

4.4.1 Model Selection and Averaging with RJMCMC

In this study, the output of the RJMCMC algorithm consists of a posterior distribution across the space of ARMA(p,q) models and their corresponding parameters. Each draw from the posterior distribution consists of information on p and q, as well as the (inverse) partial autocorrelations and consequently parameter values, and lastly the standard deviation of the disturbance corresponding to this draw. To analyze the output, two options present themselves to the researcher with respect to model choice:

1. Pick the model with the highest posterior probability
2. Average across models

Option 1 will feel more familiar to most researchers. It simply involves counting the number of draws for each combination of p and q and picking the one with the highest number of draws. It is akin to a likelihood ratio test or choosing a model based on information criteria like the Akaike Information Criterion. While one can then account for the *parameter uncertainty* conditional on the model there is no consistent way to include *model uncertainty* in the analysis of the results as one specific model is chosen. In a case where the estimates for measures of interest like the persistence measure discussed below are quite different depending on the model chosen, a phenomenon mentioned e.g. by Campbell and Mankiw (1989) for the persistence measure for France, it seems prudent to incorporate model uncertainty in the analysis.

This can be easily accomplished using the full posterior provided by RJMCMC instead of just posteriors conditional on some model chosen. Model uncertainty is accounted for by calculating the measure of interest for all draws from the posterior spanning the different models and then analyzing the resulting distribution. This approach may very well lead to wider credible sets, but this widening would then be a desirable feature as narrower sets can lead the researcher to a false sense of confidence in the results. Indeed, in quite a few cases examined here, especially when using the HP filter, considerable posterior model uncertainty remains. The results presented here account for this uncertainty.

4.5 Frequentist Regressions

The frequentist, or classical, maximum likelihood estimates are obtained using the Econometrics Toolbox of Matlab 2015a. For the frequentist estimates, the model space was constrained to include only models with autoregressive and moving average lag polynomials up to degree five.[2]

In order to pick a model, three information criteria were employed: The Akaike Information Criterion (AIC), see Akaike (1974), the Corrected Akaike Information Criterion (AICC), see Sugiura (1978) and Hurvich and Tsai (1989), and the Schwarz or Bayesian Information Criterion (BIC), see Schwarz (1978). These are given by:

$$\mathrm{AIC} = 2k - 2\ln(\hat{\mathcal{L}}), \quad \mathrm{AICC} = AIC + \frac{2k(k+1)}{n-k-1}, \quad \mathrm{BIC} = -2\ln(\hat{\mathcal{L}}) + k\ln(n)$$

with k being the number of model parameters and n the number of observations. $\hat{\mathcal{L}}$ denotes the maximized likelihood value of a model, i.e., for given ARMA orders p and q. The model chosen is then the one with the lowest value of the information criterion which is being applied.

Interestingly, the models chosen by the BIC generally exhibit impulse responses very similar to the mean and mode impulse responses obtained from RJMCMC. The AIC and AICC, on the other hand, select identical models that tend to be of higher order and the implied impulse responses differ significantly from those estimated using the other approaches.[3]

4.6 Data

Seasonally adjusted quarterly real GDP and population data used for the first experiment are taken from the OECD.Stat database. The time series for quarterly real GDP are the VOBARSA measures in this database for the period 1960:1 to 2007:4, thus excluding the Great Recession. Per capita numbers were calculated using population data from the same source.

For estimation, first differences of the natural logarithm of GDP per capita, logarithmic deviations from OLS-detrended GDP and logarithmic deviations from an Hodrick-Prescott

[2]Many authors restrict the model space even further, e.g. to $p + q \leq 6$ as e.g. in Diebold and Rudebusch (1989). The truncation of the model space chosen here is the same as in Perron (1993).

[3]Given the different philosophies behind the information criteria, this result is not surprising. See chapter 5 for a discussion of this phenomenon.

(HP) trend with the smoothing parameter λ set to 1600 were employed. The natural logarithm of GDP per capita is thus taken to be either difference stationary with drift, trend stationary with a linear trend in logarithms, or fluctuating around a logarithmic HP trend. All log growth rates and log deviations were multiplied by 100.

Since the focus of this study is the persistence of changes in GDP, the drift parameter μ for the difference stationary case is not of central interest. Thus, the first differenced series was demeaned and the remaining fluctuations taken to follow stationary and invertible zero-mean ARMA processes of unknown order. The same assumption was maintained in the estimation for the other detrending methods, i.e. the detrended data was assumed to be stationary. The drift parameter μ can be inferred from the mean in the data c together with the autoregressive coefficients for each model (or sample from the posterior) from

$$\mu = c(1 + P_1 + P_2 + \ldots + P_p)$$

.

4.7 Sampler Settings

For each of the series 4.000.000 samples from the posterior are obtained, discarding the first 1.000.000 as burn-in. The prior structure applied here assumes a priori independence for the parameters. The priors reported in Table 4.1 are the same for all variants considered.

Object	**Prior**
p	$DU(0, 10)$
q	$DU(0, 10)$
(Inverse) Partial Autocorrelation	$U(-1, 1)$
σ_ϵ	$IG(1, 1)$

Table 4.1: Prior distributions

In Table 4.1, $DU(a, b)$ denotes the discrete uniform distribution on the interval $[a, b]$, $U(c, d)$ is the continuous uniform distribution on the open interval (c, d), and $IG(e, f)$ denotes the inverse gamma distribution truncated at zero with parameters (e, f). It should be noted that, even though the prior on the orders of the lag polynomials is uniform, the proper prior on the (inverse) partial autocorrelations induces an exponentially decaying prior. If one were to increase, for example, the order of the autoregressive lag polynomial by one and set the corresponding parameter equal to zero, the likelihood would not be changed. However, the new parameter has a prior probability < 1 at all values and the posterior probability will be lowered. In this sense, additional parameters are penalized and the prior behaves implicitly like an exponential prior over $(p + q)$ which is shown in Figure 4.1. Further discussion of this feature can be found in chapter 2.

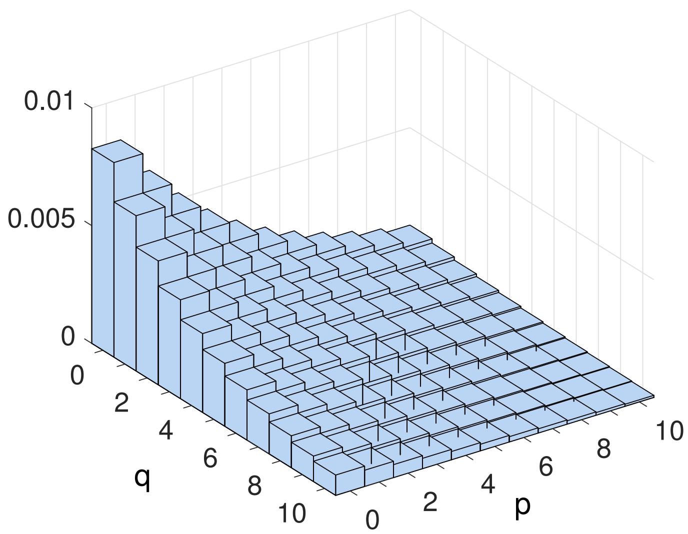

Figure 4.1: Implied posterior for model indicators

At each iteration of the RJMCMC algorithm, a new state, consisting of the model indicators p and q as well as the corresponding parameters, has to be proposed from some proposal distributions. The proposal distribution parameters were tuned by tweaking the standard deviations of the proposal distributions using short pilot runs for each of the experiments. The parameters were left constant across countries. The pilot tuning targeted acceptance rates around 20 - 30% for within-model moves, roughly in line with recommendations for fixed-dimensional random walk samplers (see, e.g. An and Schorfheide (2007)), and around 4-5% for between-model moves. This goal was not achieved in all cases. The resulting parameter values and the proposal distributions employed are reported in Table 4.2.

Detrending	Object	Proposal
First Differences	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.1^2)$
	σ_ϵ	$TN(\mu, 0.05^2)$
Linear Trend	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.03^2)$
	σ_ϵ	$TN(\mu, 0.05^2)$
HP Filter	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.025^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.07^2)$
	σ_ϵ	$TN(\mu, 0.04^2)$

Table 4.2: Proposal distributions

In Table 4.2, $DL(\mu, b)$ denotes the discretized Laplace distribution, with location parameter, μ, and shape parameter, b, such that

$$P(x|\mu) \propto exp(-b|\mu - x|) \text{ with } \mu, x \in [0, 1, \ldots, 10] \tag{4.4}$$

$TN(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2 truncated to the interval $(-1, 1)$ for the partial autocorrelations and $(0, 1000)$ for the standard deviation of the error term. Proposals are always centered around the current value of the parameter of interest as in chapter 2.

The resulting acceptance rates are presented in Table 4.3. Here, α stands for the overall acceptance rate, α_w for the acceptance rate for within model moves and α_b is the acceptance rate for between model moves. Within model moves are those, for which no change in the order of the lag polynomials is proposed. For between model moves the proposal changes least one of the two polynomial orders. The acceptance rates seem satisfactory and roughly in line with the ones in Brooks and Ehlers (2004), with the acceptance rates decreasing as the model orders increase. Even though some of the acceptance rates are low, the high number of samples used for the analysis should be sufficient to alleviate this possible problem.

Filter	FDIFF			HP			LINEAR		
	α	α_w	α_b	α	α_w	α_b	α	α_w	α_b
Canada	0.29	0.38	0.09	0.17	0.26	0.02	0.20	0.28	0.04
France	0.14	0.19	0.05	0.23	0.33	0.04	0.12	0.18	0.02
Italy	0.26	0.34	0.09	0.22	0.32	0.04	0.15	0.22	0.03
Japan	0.12	0.16	0.04	0.25	0.37	0.04	0.06	0.08	0.01
UK	0.49	0.61	0.12	0.20	0.30	0.03	0.36	0.49	0.07
US	0.22	0.29	0.07	0.19	0.28	0.03	0.23	0.33	0.05

Table 4.3: Acceptance rates for different filters
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

4.8 Impulse Responses

The following point estimates for impulse response functions at each horizon are readily available:

1. The median of the impulse response at each horizon
2. The mean of the impulse response at each horizon

Note that these estimates are different from those obtained when picking one particular model. For each draw from the posterior, the impulse response is readily calculated. In order to then, for example, obtain the median of the impulse response at some horizon, the resulting posterior distribution of the response across models and parameters at this horizon is utilized. Bayesian credible set for the responses can easily be constructed. Here, the 90% credible sets will be reported.

Together with means, medians, and credible sets, the impulse responses implied by the estimates using the information criteria will be presented and compared. All impulse responses presented are responses to a one standard deviation shock as estimated for each sample and the models from the frequentist regressions respectively.[4]

4.9 A Measure of Persistence

The measure of persistence on which this study will focus is the sum of coefficients of the infinite moving average representation of the stationary processes giving an estimate of the total persistence of the process as employed by Diebold and Rudebusch (1989) and Campbell and Mankiw (1987), among others. This measure has different interpretations, depending on the nature of the underlying model.

Let $C(L)$ denote the infinite order polynomial in the lag operator given by the infinite moving average representation of a stationary ARMA(p,q) model and let $C_n(1)$ be the sum

[4] This is necessary as the unconditional variance of an ARMA model is a function of not only the standard deviation of the disturbance, but also the AR and MA polynomials. If the model or its parameters values change, the corresponding standard deviation has to change as well to match the variation in the data.

of the first n coefficients:

$$P(L)y_t = Q(L)\epsilon_t \tag{4.5}$$

$$y_t = \frac{Q(L)}{P(L)}\epsilon_t = C(L)\epsilon_t = (1 + C_1 L + C_2 L^2 + \ldots)\epsilon_t \tag{4.6}$$

$$C_n(1) = 1 + \sum_{i=1}^{n} C_i \tag{4.7}$$

$C_n(1)$ thus gives the *cumulated* response to a shock up to horizon n.

What information does this statistic convey? Consider first a model in which the y_t are first-differenced log GDP per capita data points. In this setup, C_i gives the effect of a disturbance on the *growth rate* occurring at time t on the *growth rate* at time $t+i$. The cumulative effect on the *level* of GDP at time $t+n$ is then given by $C_n(1)$. $C_n(1)$ is thus the change in one's forecast for the level of GDP at time $t+n$ after observing a unit shock in t. For a random walk holds, for example, $C_n(1) = 1 \forall n$, while if the series were trend stationary, $C_n(1)$ would converge to zero with increasing n as the effect of the shock on the level of GDP vanishes with trend-reversion (see Campbell and Mankiw (1987) for further discussion).

In a trend-stationary world, be it a Hodrick-Prescott or a linear trend, the measure will give the undiscounted sum of departures from the trend in future periods in log points. The higher $C_n(1)$, the more pronounced the departure of GDP from its trend up to time $t+n$ after a shock occurring in period t.

4.10 Kolmogorov-Smirnov Test

In order to compare the estimates from RJMCMC output across countries– apart from optical inspection of the impulse responses and posterior distributions for the statistics considered and corresponding intracranial trauma tests– a more formal means of comparison will be employed here. Since RJMCMC delivers a posterior distribution for the persistence measures, I can test whether any two sets of samples from the posteriors seem to be generated by the same distribution.

The test employed here is the two-sample Kolmogorov-Smirnov test, which has equality of the distributions in the two samples as its null hypothesis. The corresponding test statistic for two distributions a and b is given by

$$KSS_{a,b} = \sup_x |F_a(x) - F_b(x)|$$

where $F_a(x)$ and $F_b(x)$ denote the cumulative distribution functions associated with the distributions a and b. The critical values for this statistic are given by

$$KSS_\alpha = c(\alpha)\sqrt{\frac{n_a + n_b}{n_a n_b}}$$

where n_a and n_b are the sample sizes for posteriors a and b respectively and $c(\alpha)$ is a coefficient depending on the chosen significance level α:

α	0.05	0.01
$c(\alpha)$	1.36	1.63

4.11 Results

In the following sections, the results of the estimation using GDP growth rates and the robustness checks using Hodrick-Prescott as well as OLS linear detrending will be presented.

4.11.1 GDP Growth Rates

This section presents the results obtained using first differences of the natural logarithm of GDP. It is thus primarily concerned with the dynamics of GDP growth rates.

Model Choice One of the main advantages of RJMCMC is the possibility to plot and inspect the posterior distribution *across* models. Here, the role of the model indicator is played by the orders of the two lag polynomials, p for the AR polynomial and q for the MA polynomial. The pair (p, q) then identifies one model. Figure 4.2 shows the posteriors over the model indicators p and q for all six countries.

Inspection of the plots shows clear differences in the posteriors over the models for the six countries. Notably, for the UK the pure random walk model for the level of GDP is clearly preferred by RJMCMC. There are very few samples with low order AR and MA polynomials. This result will be revisited later.

In contrast, the posterior for France has the most posterior mass assigned to the ARMA(3,1) model with quite substantial posterior uncertainty regarding the model and the possibility of multi-modality with the second mode at the ARMA(1,2) model. The posterior for Japan has its mode at the ARMA(2,2) model with similarly pronounced posterior model uncertainty. These higher-order and mixed models allow for more intricate and possibly more persistent impulse responses as will become obvious in the next section.

The posterior mode in the (p, q) space for the US is at the ARMA(2,0) model, a result in line with the ones presented in chapter 2 with the rest of the posterior mass clustered around this point. The posterior for Canada exhibits a similar picture but clearly favors a simple AR(1) model over the AR(2) specification preferred for the US. Both posteriors also show strong similarities with the one for Italy. The posterior for Italy is, however, more dispersed around the mode at the AR(1) model with almost negligible differences in posterior probabilities for the neighboring models ARMA(1,1) and AR(2), indicating higher model uncertainty compared to e.g. Canada for which the posterior distribution has a much more pronounced mode at the AR(1) model. It should be noted, that the AR(1) model imposes significant restrictions on the shape of the impulse response function as an AR(1) model will always exhibit exponential decay of the impulse response, oscillating or not.

Thus, the posteriors over the model indicators already hint at differing dynamic behavior of the GDP growth rates across countries.

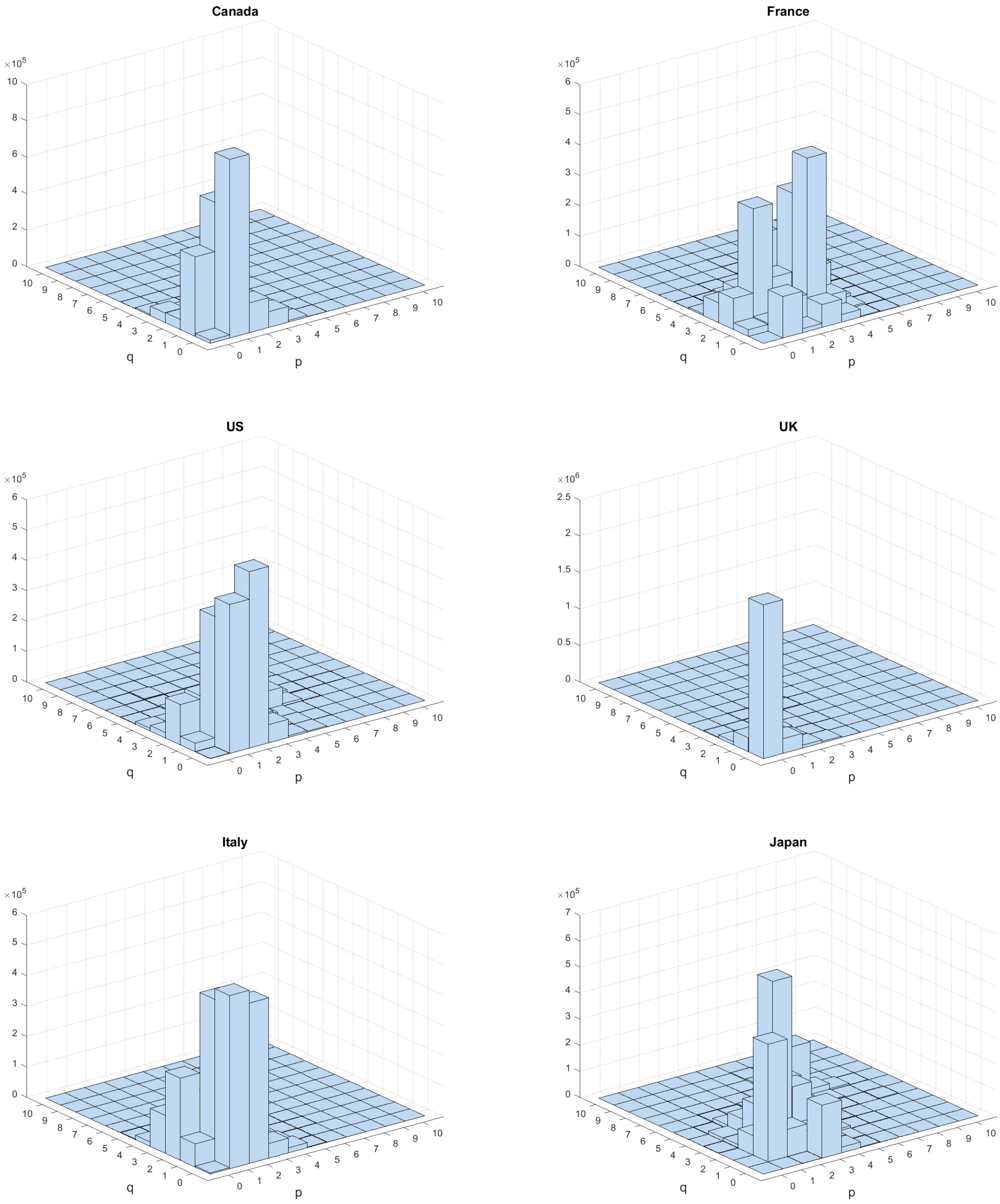

Figure 4.2: Posterior distributions for model indicators for first differences

Impulse Responses I now turn to an analysis of the estimated impulse responses to a positive one standard deviation shock to the growth rate for the six countries. The impulse responses are presented in Figure 4.3. The impulse responses and the persistence measures are calculated using every 30th draw from the posterior giving 1.000.000 draws to keep computation time manageable. This approach, called thinning, also reduces the autocorrelations in the samples from the posterior which is very much desirable as inference is based on the assumption that the samples are independently distributed. The models on which the frequentist impulse responses are based are presented in Table 4.4.

A few observations can be made from visual inspection of the plots. Models chosen by the AIC and AICC criteria coincide among the two criteria for all six countries and the models chosen by the BIC are significantly closer to the means and modes of the impulse responses from RJMCMC with BIC and RJMCMC choosing more parsimonious models. AIC and AICC choose models characterized by higher order lag polynomials as well as complex conjugate roots in the AR-polynomials, as evident in the dampened oscillations in the impulse responses. The means and medians of the impulse responses are similar to one another. With the exception of Japan and to some degree, France, the credible sets for the impulse responses are relatively tight despite model uncertainty present in the posterior.

Turning to the differences between countries, the response of US, Italian and Canadian growth rates to a shock show a similar pattern of persistence: the mean and median responses decay geometrically until reaching zero at a horizon of about 6 quarters. Notably, the credible sets for the US compared to the ones for Canada and Italy are somewhat different. The credible set for the former is wider, includes responses below zero, and the lower bound remains below zero up to 30 quarters. The credible sets for the impulse responses for the two latter countries do not encompass negative responses at any horizon and the upper bound reaches zero after 20 quarters and 17 quarters, for Canada and Italy respectively.

The impulse response for the UK reflects the large posterior mass put on the pure random walk (ARIMA(0,1,0)) by RJMCMC. The credible sets allow for some very limited persistence

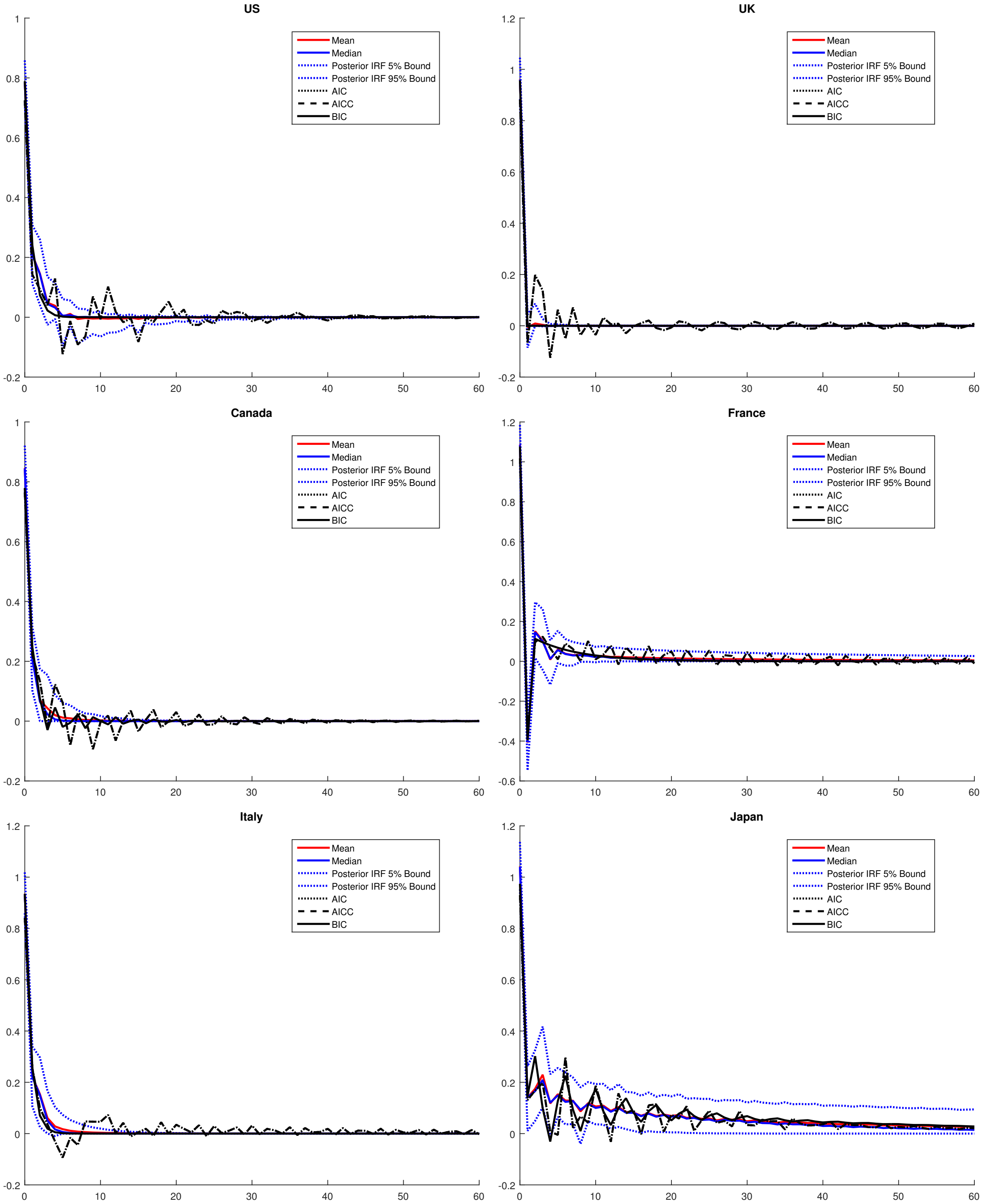

Figure 4.3: Estimated impulse responses for first differences

Country	Criterion	P_1	P_2	P_3	P_4	P_5	Q_1	Q_2	Q_3	Q_4	Q_5	σ_e
Canada	AIC	0.717	−0.273	0.747	−0.592		−0.428	0.241	−0.816	0.601		0.764
		(0.150)	(0.066)	(0.053)	(0.117)		(0.161)	(0.095)	(0.097)	(0.136)		(0.063)
	AICC	0.717	−0.273	0.747	−0.592		−0.428	0.241	−0.816	0.601		0.764
		(0.150)	(0.066)	(0.053)	(0.117)		(0.161)	(0.095)	(0.097)	(0.136)		(0.063)
	BIC	−1.002	−0.417	0.191			1.320	0.832				0.779
		(0.065)	(0.079)	(0.057)			(0.054)	(0.057)				(0.056)
France	AIC	−1.016	−0.085	1.000	0.560	0.195	0.651	−0.192	−0.813			1.017
		(0.108)	(0.198)	(0.117)	(0.099)	(0.083)	(0.116)	(0.155)	(0.122)			(0.097)
	AICC	−1.016	−0.085	1.000	0.560	0.195	0.651	−0.192	−0.813			1.017
		(0.108)	(0.198)	(0.117)	(0.099)	(0.083)	(0.116)	(0.155)	(0.122)			(0.097)
	BIC	0.845					−1.212	0.415				1.077
		(0.091)					(0.085)	(0.038)				(0.059)
Italy	AIC	0.625	0.922	−0.469	−0.753	0.639	−0.322	−0.992	0.158	0.722	−0.482	0.841
		(0.212)	(0.117)	(0.225)	(0.149)	(0.127)	(0.230)	(0.181)	(0.227)	(0.203)	(0.152)	(0.058)
	AICC	0.625	0.922	−0.469	−0.753	0.639	−0.322	−0.992	0.158	0.722	−0.482	0.841
		(0.212)	(0.117)	(0.225)	(0.149)	(0.127)	(0.230)	(0.181)	(0.227)	(0.203)	(0.152)	(0.058)
	BIC	0.273										0.934
		(0.057)										(0.057)
Japan	AIC	−0.866	−0.209	0.348	0.783	0.731	1.007	0.511	0.050	−0.602	−0.858	0.930
		(0.031)	(0.036)	(0.036)	(0.055)	(0.034)	(0.061)	(0.095)	(0.093)	(0.088)	(0.053)	(0.084)
	AICC	−0.866	−0.209	0.348	0.783	0.731	1.007	0.511	0.050	−0.602	−0.858	0.930
		(0.031)	(0.036)	(0.036)	(0.055)	(0.034)	(0.061)	(0.095)	(0.093)	(0.088)	(0.053)	(0.084)
	BIC	0.973	−0.801	0.783			−0.823	0.965	−0.863			0.974
		(0.032)	(0.026)	(0.025)			(0.059)	(0.031)	(0.059)			(0.090)
UK	AIC	−0.235	−0.544	−0.751			0.157	0.750	0.918	−0.041	0.292	0.882
		(0.095)	(0.047)	(0.092)			(0.106)	(0.073)	(0.101)	(0.059)	(0.060)	(0.071)
	AICC	−0.235	−0.544	−0.751			0.157	0.750	0.918	−0.041	0.292	0.882
		(0.095)	(0.047)	(0.092)			(0.106)	(0.073)	(0.101)	(0.059)	(0.060)	(0.071)
	BIC											0.957
												(0.054)
US	AIC	−0.140	0.343	−0.169	−0.726		0.336	−0.187	0.175	0.901		0.725
		(0.067)	(0.049)	(0.060)	(0.043)		(0.064)	(0.068)	(0.067)	(0.052)		(0.042)
	AICC	−0.140	0.343	−0.169	−0.726		0.336	−0.187	0.175	0.901		0.725
		(0.067)	(0.049)	(0.060)	(0.043)		(0.064)	(0.068)	(0.067)	(0.052)		(0.042)
	BIC	0.305										0.789
		(0.063)										(0.047)

Table 4.4: Frequentist regression results for first differences

due to the few samples with low-order ARMA model in the posterior, but collapse completely after about 6 quarters. Negative responses are included in the credible set at a horizon of 1 quarter.

The impulse response functions for France exhibit particularly interesting dynamics. A shock to the growth rate of real GDP leads to a strongly *negative* response of the growth rate one quarter after the shock with a magnitude of about 40% of one standard deviation of the disturbance, thereafter turning positive again. The credible sets do not even allow for a zero or positive response after one quarter. In quarter two after the shock, the mean response, the median response, and the credible sets are all positive. In the third quarter following the shock the credible sets allow for a negative response once more. This shape is also present in the impulse responses based on frequentist estimates. However, the AIC and AICC pick models with strongly and very persistent oscillatory behavior. The credible sets for France include positive responses at horizons as long as 60 quarters at which point the oscillations from the two aforementioned information criteria are still present.

Equally interesting is the impulse response for Japan, for which the means and medians exhibit a slightly oscillatory pattern. The response always remains positive. The credible sets for Japan are considerably wider than those for the other countries and the response is very persistent, with the mean response being 0.03 log points after 40 quarters and the credible set encompassing the area between zero and 0.116 log points. All information criteria pick models with strongly oscillatory behavior. Interestingly, these results fit squarely with narratives about the French and Japanese economy being slow to adjust to shocks.

To conclude, from the perspective of impulse response functions, the dynamic behavior of GDP growth rates seems to differ quite strongly between the countries studied with the greatest similarities among US, Canada, and Italy.

Persistence I now turn to the discussion of estimates for the persistence measure $C_n(1)$ at different horizons. Figures 4.4 and 4.5 present posterior distributions for $C_n(1)$ for horizons of

20, 40 and 60 quarters.

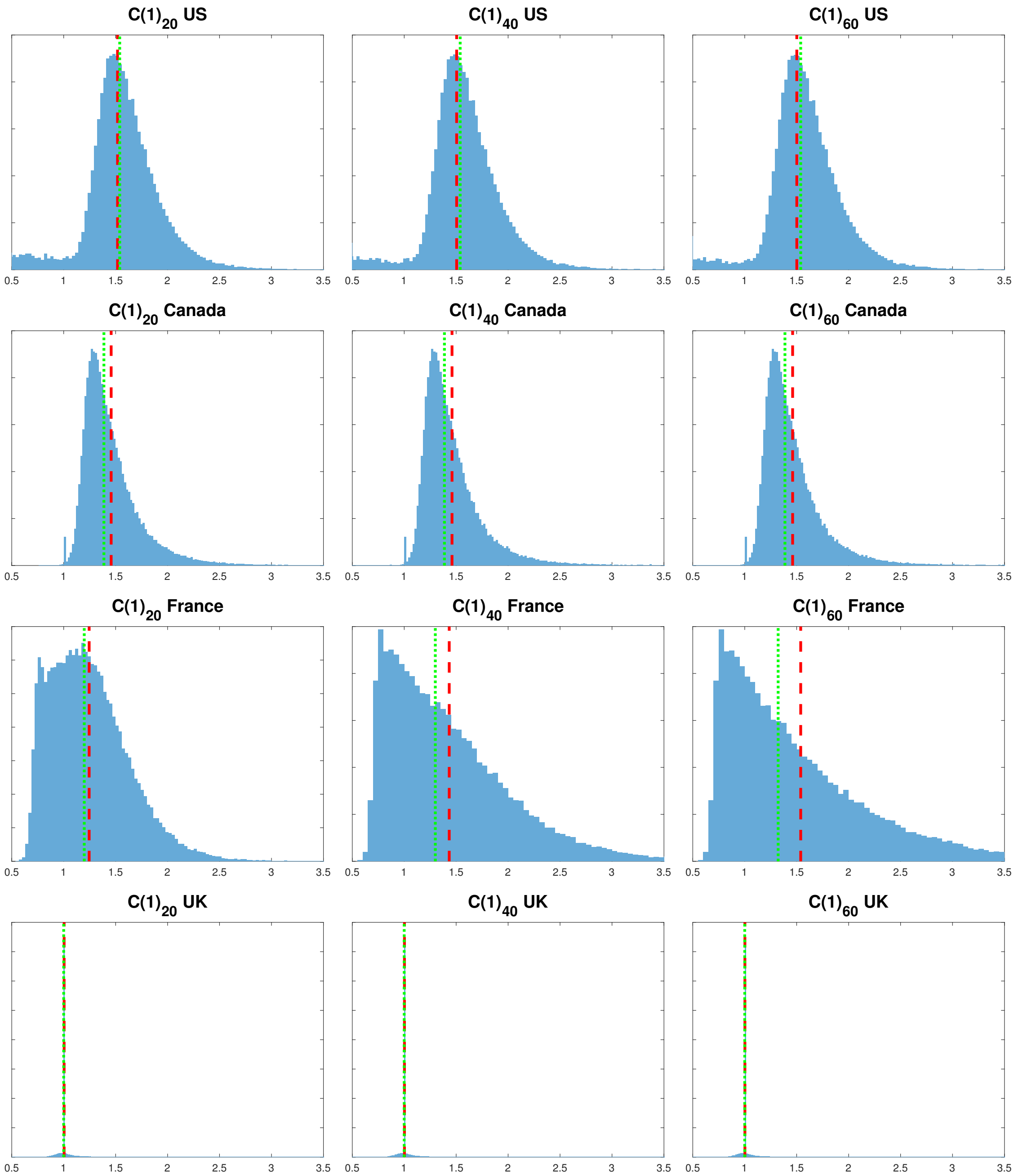

Figure 4.4: $C_n(1)$ for first differences:
Mean: Dashed line; Median: Dotted line

Tables 4.5 and 4.6 report point estimates for the persistence measure at different horizons

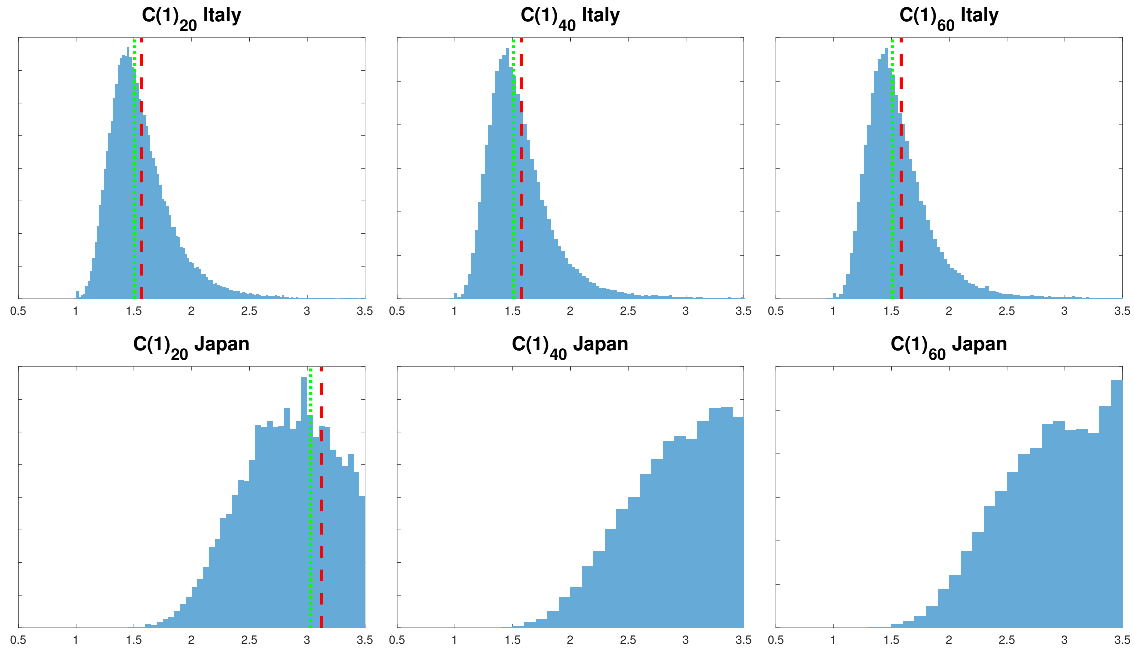

Figure 4.5: $C_n(1)$ for first differences:
Mean: Dashed line; Median: Dotted line

from RJMCMC and frequentist methods respectively. The table for the RJMCMC results contains the posterior mean and [median] as well as the 90% credible sets in the second row.

Inspection of the posteriors again reveals differences similar to those observed in the impulse response functions. The posterior at all horizons for the UK has its mode at $C(1)_n = 1$ with very little variation, which is to be expected given the foregoing analysis since the clearly preferred model for the UK is a pure random walk. Not surprisingly, the dispersion of the posterior distributions mirrors the width of the credible sets in the impulse responses. For the US, Italy, France, and Canada the posterior distributions have means clustered around 1.5 at a horizon of 60 quarters, with a range of 1.46 for Canada to 1.58 for Italy. The shapes and variances of the posteriors also differ between these countries. Additionally, the medians and means of the posteriors appear stable across horizons for all countries except France and Japan.

The behavior of the posterior mean and median responses is different for these two countries, for which the posterior distributions shift to the right as the horizon increases. This phenomenon is most pronounced for Japan. This higher persistence is already visible in the impulse response functions: the very persistent impulse response implies that the growth rate will be above its average for a longer period following a positive shock with the resulting effect on the level of GDP accumulating more strongly over time. The shape of the posterior distribution for Japan changes slightly across horizons with the lower bound increasing until a horizon of 40 quarters, after which only the upper bound increases further. As a result, both mean and median tend to grow and the credible sets widen as the horizon increases.

For France, the behaviors of the mean and median are different. While the mean grows as the horizon increases from 40 to 60 quarters, the median remains roughly constant, due to an increase of the upper bound of the credible set while the lower bound is constant. The change in the shape of the posterior is clearly visible in Figure 4.4.

Horizon	5	10	20	40	60
Canada	1.42 [1.38]	1.45 [1.39]	1.46 [1.39]	1.46 [1.39]	1.46 [1.39]
	[1.16; 1.8]	[1.16; 1.94]	[1.16; 2]	[1.16; 2.01]	[1.16; 2.01]
France	0.921 [0.913]	1.07 [1.05]	1.25 [1.2]	1.43 [1.3]	1.54 [1.32]
	[0.726; 1.15]	[0.744; 1.45]	[0.746; 1.93]	[0.746; 2.6]	[0.746; 3.07]
Italy	1.51 [1.49]	1.55 [1.5]	1.56 [1.51]	1.58 [1.51]	1.58 [1.51]
	[1.22; 1.89]	[1.22; 2.03]	[1.22; 2.1]	[1.22; 2.14]	[1.22; 2.14]
Japan	1.78 [1.75]	2.33 [2.28]	3.12 [3.03]	4.11 [3.92]	4.72 [4.38]
	[1.42; 2.23]	[1.77; 3.06]	[2.2; 4.35]	[2.39; 6.48]	[2.4; 8.22]
UK	1.01 [1]	1.01 [1]	1 [1]	1 [1]	1 [1]
	[0.92; 1.12]	[0.917; 1.13]	[0.914; 1.13]	[0.914; 1.13]	[0.914; 1.13]
US	1.56 [1.54]	1.56 [1.54]	1.52 [1.54]	1.5 [1.54]	1.5 [1.54]
	[1.22; 1.98]	[0.962; 2.11]	[0.557; 2.14]	[0.368; 2.14]	[0.31; 2.14]

Table 4.5: $C_n(1)$ for first differences at different horizons; RJMCMC estimates
Mean and [median] with [90% credible sets] in the second row

Turning to the frequentist estimates in Table 4.6, the differences in the behavior of the

Horizon	5	10	20	40	60
Canada	1.68; 1.41	1.51; 1.42	1.51; 1.42	1.49; 1.41	1.49; 1.41
France	0.925; 0.964	1.18; 1.16	1.52; 1.28	1.75; 1.31	1.83; 1.31
Italy	1.32; 1.38	1.41; 1.38	1.65; 1.38	1.93; 1.38	2.09; 1.38
Japan	1.54; 1.66	2.2; 2.29	2.94; 3.12	3.97; 4.36	4.53; 5.1
UK	1.23; 1	1.18; 1	1.2; 1	1.21; 1	1.22; 1
US	1.39; 1.44	1.24; 1.44	1.34; 1.44	1.31; 1.44	1.31; 1.44

Table 4.6: $C_n(1)$ for first differences at different horizons
Frequentist estimates for AIC; BIC

point estimates between countries are clearly visible again. The frequentist estimates appear mostly consistent with the estimates from RJMCMC even though the models chosen differ significantly, especially for the AIC and AICC.[5] The clustering of the estimates at longer horizons is present in those based on the BIC, but not in those using the AIC. The frequentist estimates are contained in the credible sets with the exception of the AIC estimate for the UK.

It is instructive to compare these estimates to the results of Campbell and Mankiw (1989) (henceforth CM) who use quarterly real *GNP* for the G7 to estimate $C_n(1)$. Their results can be found in Table 4.7 together with means and [medians] from RJMCMC. The pattern of an increase in $C_n(1)$ as n increases is present for all countries in their results in contrast to the findings presented here. There is no clear pattern regarding the relative size of the estimates from CM and RJMCMC.

Table 4.8 presents a ranking of the six countries based on the estimated $C_n(1)$ with the first-ranking country being the most persistent. Clearly, the pattern of persistence across countries leads to a similar persistence ranking for all estimates with the exception of the US being ranked consistently lower by CM and AIC. The ranking using the BIC and medians coincide well. Also, the ranking appears stable for each method when changing the horizon.

It should be noted, however, that for countries for which the estimates are close, the respective values lie well within the 90% credible sets of one another. For example, the mean

[5] As the estimates using AICC and AIC are identical, only the AIC estimates are presented here and below.

Horizon:	20	40	60
Canada	1.57	1.88	1.92
	1.46 [1.39]	1.46 [1.39]	1.46 [1.39]
France	1.39	1.86	2.06
	1.25 [1.2]	1.43 [1.3]	1.54 [1.32]
Italy	1.44	1.96	2.45
	1.56 [1.51]	1.58 [1.51]	1.58 [1.51]
Japan	2.31	3.18	3.71
	3.12 [3.03]	4.11 [3.92]	4.72 [4.38]
UK	0.76	0.88	0.94
	1 [1]	1 [1]	1 [1]
US	1.21	1.22	1.25
	1.52 [1.54]	1.5 [1.54]	1.5 [1.54]

Table 4.7: $C_n(1)$: Results for first differences from CM in the first row, posterior mean and [median] in the second

Horizon	20					40					60				
Estimate	Mean	Median	CM	AIC	BIC	Mean	Median	CM	AIC	BIC	Mean	Median	CM	AIC	BIC
Canada	4	4	2	4	4	4	4	3	4	3	5	4	4	4	3
France	5	5	4	3	5	5	5	4	3	5	3	5	3	3	5
Italy	2	3	3	2	3	2	3	2	2	4	2	3	2	2	4
Japan	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
UK	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
US	3	2	5	5	2	3	2	5	5	2	4	2	5	5	2

Table 4.8: Ranking by persistence for first differences

of $C_n(1)$ for Italy at a horizon of 40 quarters is equal to 1.58 with a credible set in $[1.22; 2.14]$. The credible set thus contains the point estimates for Canada, France, and the US.

Kolmogorov-Smirnov Test for C(1) In order to gain a more complete picture regarding the differences in the persistence estimates, this section compares the whole posterior distributions for $C_n(1)$ at different horizons. Table 4.9 presents the test statistic for a horizon of 40 quarters. Results for horizons 5, 10, 20, 30, 50, and 60 can be found in the appendix. All pairwise two-sample Kolmogorov-Smirnov test applied to the posteriors reject the null hypothesis at the 1% level. Interestingly, according to the test statistic at different horizons, the posteriors for the US and Italian economies are most similar. Furthermore, the US, Canada and Italy form a trio with fairly similar posterior distributions of $C_n(1)$ at all horizons compared to the other countries.

	Canada	France	Italy	Japan	UK	US
Canada	0	0.35207 (*)	0.2237 (*)	0.94847 (*)	0.92226 (*)	0.24055 (*)
France	0.35207 (*)	0	0.39423 (*)	0.87453 (*)	0.58686 (*)	0.32605 (*)
Italy	0.2237 (*)	0.39423 (*)	0	0.93387 (*)	0.94124 (*)	0.09863 (*)
Japan	0.94847 (*)	0.87453 (*)	0.93387 (*)	0	0.99633 (*)	0.93799 (*)
UK	0.92226 (*)	0.58686 (*)	0.94124 (*)	0.99633 (*)	0	0.84812 (*)
US	0.24055 (*)	0.32605 (*)	0.09863 (*)	0.93799 (*)	0.84812 (*)	0

Table 4.9: K-S test for $C(1)_{40}$ for first differences

Summary To summarize, while differences exist in the persistence estimates, the posteriors contain significant uncertainty. The economies of both the UK and Japan, however, exhibit a behavior that differs strongly from that seen in other countries under inspection. The results of CM are roughly in line with the results presented here, with Japan being highly persistent and the UK exhibiting the lowest degree of persistence in growth rates. The estimates using the BIC are closest to the estimates obtained with RJMCMC.

4.11.2 Robustness

Since it is well known that the detrending method chosen can have significant impact on empirical results, see e.g. Canova (1998), the results from the difference stationary perspective will now be compared with the results obtained using linearly detrended and Hodrick-Prescott filtered data.

Linear Detrending

This section investigates whether the ranking of persistence obtained taking the first-difference stationary perspective will hold up under ordinary least squares (OLS) linear detrending. RJM-CMC was applied to the logarithmic deviations of GDP from an OLS linear trend.

Model Choice Comparing the posterior distributions for the model indicators for the six countries presented in Figure 4.6, significant differences in the posteriors are immediately obvious. Notably, for the UK, the preferred AR(1) model is again the most parsimonious among the countries with very limited posterior uncertainty. Furthermore, also in the linear trend world, the posteriors for Canada and the US seem quite similar, albeit with different modes at the AR(2) model for Canada and the AR(3) model for the US where the modes were at the AR(1) and AR(2) model respectively from the difference stationary perspective.

The posterior for Italy now indicates the possibility of multi-modality. The model at the mode here is an AR(3), albeit exhibiting significant posterior uncertainty and very little difference in the posterior probability compared to the AR(2) model. This observation is again in line with the results from the analysis of growth rates where the posterior probabilities for Italy were quite close for the group of models clustered around the mode.

The posteriors for France and Japan show the greatest posterior uncertainty regarding the model with pronounced multi-modality for Japan. The preferred models for France and Japan are ARMA(4,1) and AR(4) respectively. The second mode for Japan is at the ARMA(3,2) model. France exhibits more dispersed clustering of high posterior probability models around

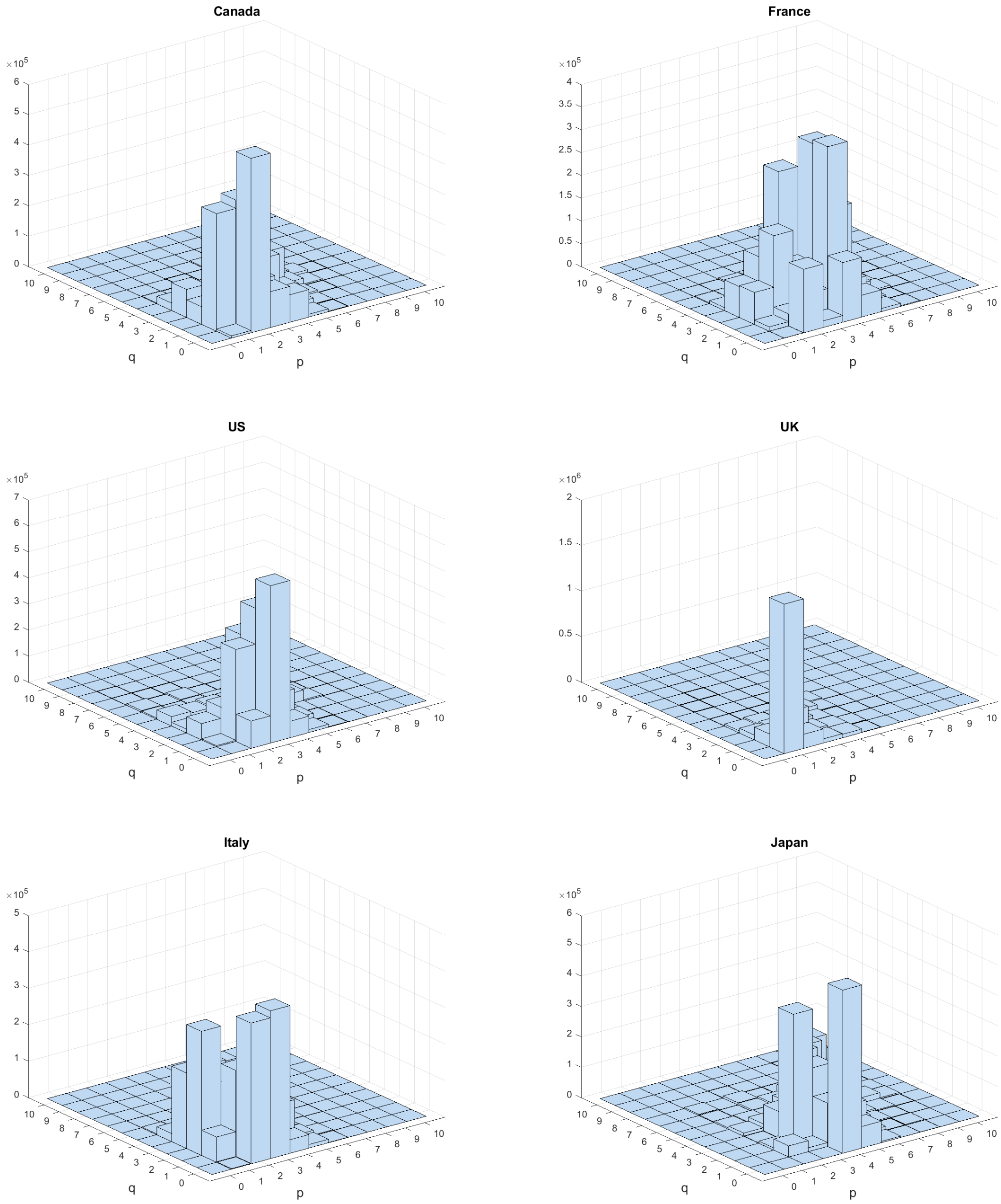

Figure 4.6: Posterior distributions for model indicators for linear trend

the mode and multi-modality is diminished compared to the growth rate case.

Impulse Responses The impulse response functions for the linear trend perspective are reported in Figure 4.7. The frequentist estimates are presented in Table 4.10. Not surprisingly, the impulse response functions show substantially more persistence and are different in shape compared to the ones obtained under first differencing.

The estimates obtained using RJMCMC compared to those using the information criteria differ more strongly. This difference is especially pronounced in the case of Italy and France, where the estimates for Italy from all three information criteria are not covered by the credible sets. For France, the impulse response implied by the model selected by the AIC and AICC lies also completely outside the credible set while the one chosen by the BIC lies within. The impulse responses for the models selected by the information criteria for the US basically trace out the lower bound of the credible set. For the UK, the model chosen by the BIC coincides with mean and median responses from RJMCMC. The frequentist impulse responses using AIC and AICC show small oscillations. These oscillations feature in the RJMCMC estimates only for Canada, and for Canada as well as Italy for BIC estimates. The choices of the three information criteria coincide in the case of Italy.

The impulse response functions for the UK do not show the familiar hump-shaped pattern, a consequence of the dominant model in the posterior being AR(1). The other countries, however, exhibit a hump-shaped response, albeit with substantially differing persistence. The mean response for the US remains slightly positive up to 60 quarters, but is already at a low level of 0.06 log points after 30 quarters and a response of zero is contained in the credible set starting in quarter 14 after the shock. In comparison, the impulse response for Canada converges to zero at a much slower rate and the credible sets do not contain a zero response even after 60 quarters. The impulse response for Italy, while hump-shaped, is even more persistent: it reaches the level of one shock standard deviation only after about 55 quarters.

Interestingly, the kink in the impulse response function for France is still present here.

Country	Criterion	P_1	P_2	P_3	P_4	P_5	Q_1	Q_2	Q_3	Q_4	Q_5	σ_e
Canada	AIC	−0.272	0.513	0.683			1.556	1.229	0.469	0.292	0.144	0.764
		(0.041)	(0.033)	(0.046)			(0.079)	(0.136)	(0.136)	(0.109)	(0.059)	(0.055)
	AICC	−0.272	0.513	0.683			1.556	1.229	0.469	0.292	0.144	0.764
		(0.041)	(0.033)	(0.046)			(0.079)	(0.136)	(0.136)	(0.109)	(0.059)	(0.055)
	BIC	−0.284	0.517	0.703			1.575	1.167	0.218			0.774
		(0.042)	(0.033)	(0.047)			(0.073)	(0.106)	(0.062)			(0.056)
France	AIC	0.092	1.195	0.897	−0.763	−0.426	0.504	−0.504	−1.000			1.010
		(0.037)	(0.025)	(0.043)	(0.023)	(0.029)	(0.064)	(0.051)	(0.058)			(0.081)
	AICC	0.092	1.195	0.897	−0.763	−0.426	0.504	−0.504	−1.000			1.010
		(0.037)	(0.025)	(0.043)	(0.023)	(0.029)	(0.064)	(0.051)	(0.058)			(0.081)
	BIC	1.882	−0.887				−1.281	0.430				1.057
		(0.066)	(0.065)				(0.057)	(0.035)				(0.059)
Italy	AIC	0.314	1.387	0.231	−0.937		0.922	−0.403	−1.000	−0.330	−0.190	0.841
		(0.011)	(0.014)	(0.011)	(0.010)		(0.071)	(0.129)	(0.104)	(0.094)	(0.110)	(0.059)
	AICC	0.314	1.387	0.231	−0.937		0.922	−0.403	−1.000	−0.330	−0.190	0.841
		(0.011)	(0.014)	(0.011)	(0.010)		(0.071)	(0.129)	(0.104)	(0.094)	(0.110)	(0.059)
	BIC	0.314	1.387	0.231	−0.937		0.922	−0.403	−1.000	−0.330	−0.190	0.841
		(0.011)	(0.014)	(0.011)	(0.010)		(0.071)	(0.129)	(0.104)	(0.094)	(0.110)	(0.059)
Japan	AIC	0.682	1.500	−0.762	−0.752	0.329	0.384	−1.001	−0.232	0.148	−0.159	0.986
		(0.278)	(0.117)	(0.436)	(0.057)	(0.234)	(0.272)	(0.240)	(0.219)	(0.208)	(0.087)	(0.091)
	AICC	0.391	1.500	−0.214	−0.735	0.054	0.712	−0.766	−0.611			0.997
		(0.159)	(0.107)	(0.201)	(0.056)	(0.101)	(0.136)	(0.077)	(0.101)			(0.087)
	BIC	1.127	0.067	0.099	−0.302							1.028
		(0.071)	(0.108)	(0.110)	(0.078)							(0.095)
UK	AIC	0.344	0.774	0.475	−0.679		0.627	−0.088	−0.565	0.200		0.888
		(0.175)	(0.079)	(0.079)	(0.148)		(0.165)	(0.263)	(0.209)	(0.078)		(0.067)
	AICC	0.344	0.774	0.475	−0.679		0.627	−0.088	−0.565	0.200		0.888
		(0.175)	(0.079)	(0.079)	(0.148)		(0.165)	(0.263)	(0.209)	(0.078)		(0.067)
	BIC	0.949										0.943
		(0.024)										(0.051)
US	AIC	1.081	0.764	−0.491	−0.895	0.522	0.195	−0.820	−0.656	0.473		0.708
		(0.241)	(0.133)	(0.221)	(0.056)	(0.166)	(0.264)	(0.195)	(0.118)	(0.226)		(0.043)
	AICC	1.081	0.764	−0.491	−0.895	0.522	0.195	−0.820	−0.656	0.473		0.708
		(0.241)	(0.133)	(0.221)	(0.056)	(0.166)	(0.264)	(0.195)	(0.118)	(0.226)		(0.043)
	BIC	1.785	−0.818				−0.615					0.748
		(0.063)	(0.059)				(0.099)					(0.041)

Table 4.10: Frequentist regression results for linear trend

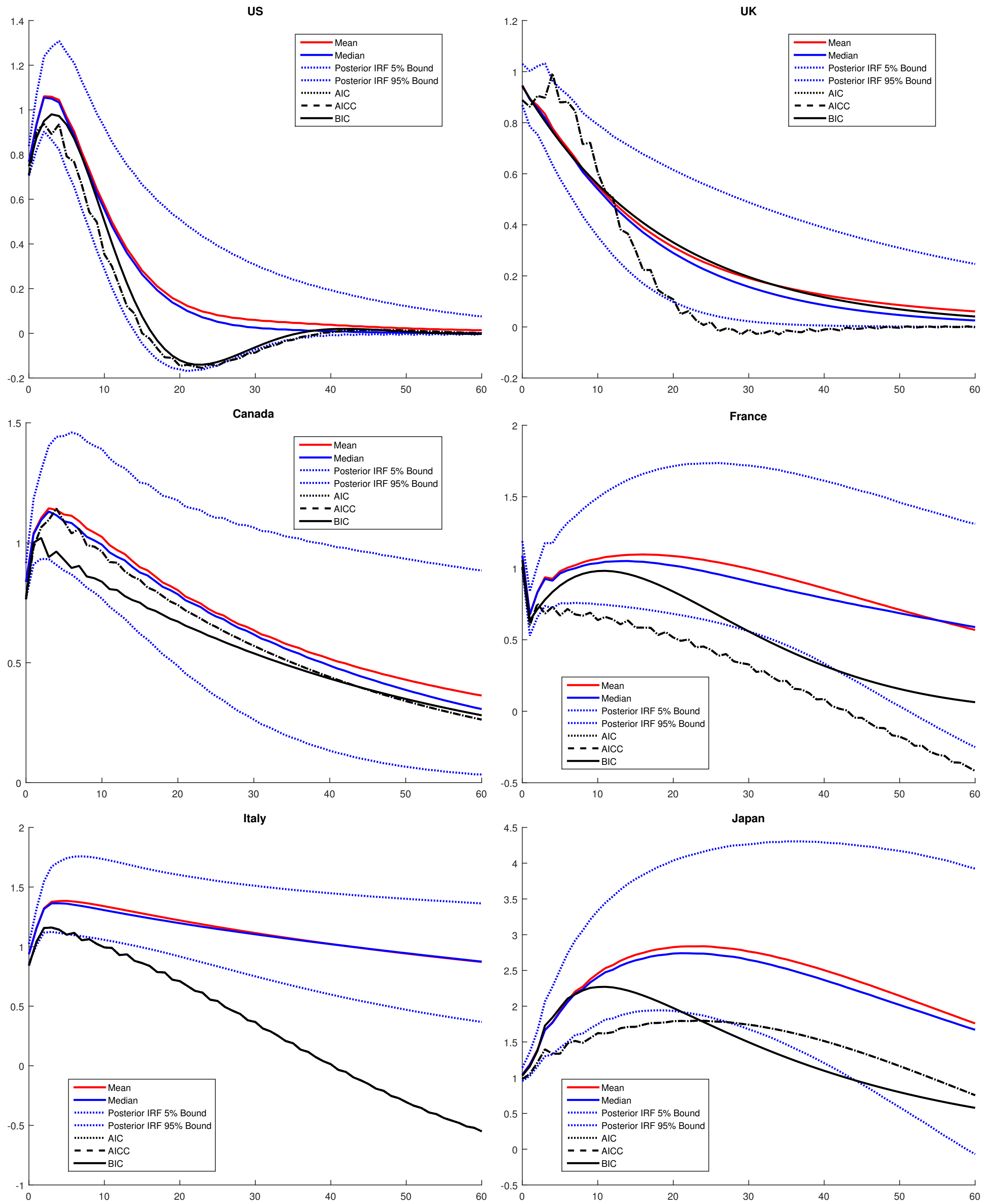

Figure 4.7: Estimated impulse responses for linear trend

After the initial reversion towards the trend, the response of GDP is hump-shaped as well. The credible sets for France are quite wide and contain a zero response after 50 quarters but the mean and median responses are still only slightly below the initial response at the time of the shock after 60 quarters.

The response for Japan is even more persistent. Mean and median response as well as the bounds of the credible sets *increase* until reaching a maximum only after about 22 quarters. The credible sets, however, are quite wide again, including a zero response after 60 quarters. For this extreme case, the models chosen by AIC and AICC exhibit a response that is less pronounced in terms of magnitude but similar in shape while the response of the model chosen by the BIC peaks already after 10 quarters.

The substantial persistence in the impulse responses found here and the higher orders of the lag polynomials of the models selected can be seen as an indication that it might be reasonable to adopt a difference stationary perspective to more parsimoniously capture the dynamics of the series. Apart from the impulse response for the US, a shock to GDP causes a significant departure from the trend even after 10 years, pointing towards substantial persistence in the response to a shock.

Persistence Figures 4.8 and 4.9 show the posterior distributions for the persistence measure for the six countries under linear detrending. It should, however, be kept in mind that the interpretation of the measure is different with linear detrending as explained in the foregoing.

The means and medians of the posterior distributions of the persistence measure move to the right as the horizon increases. The US and UK show the smallest change in $C_n(1)$ with changing horizon, as well as the lowest persistence. Japan again exhibits by far the largest persistence. The dispersion of the posterior distributions again reflects the width of the credible sets for the impulse responses.

Tables 4.11 and 4.12 present point estimates for the persistence measure from RJMCMC and the frequentist methods respectively. The RJMCMC estimates do differ more significantly

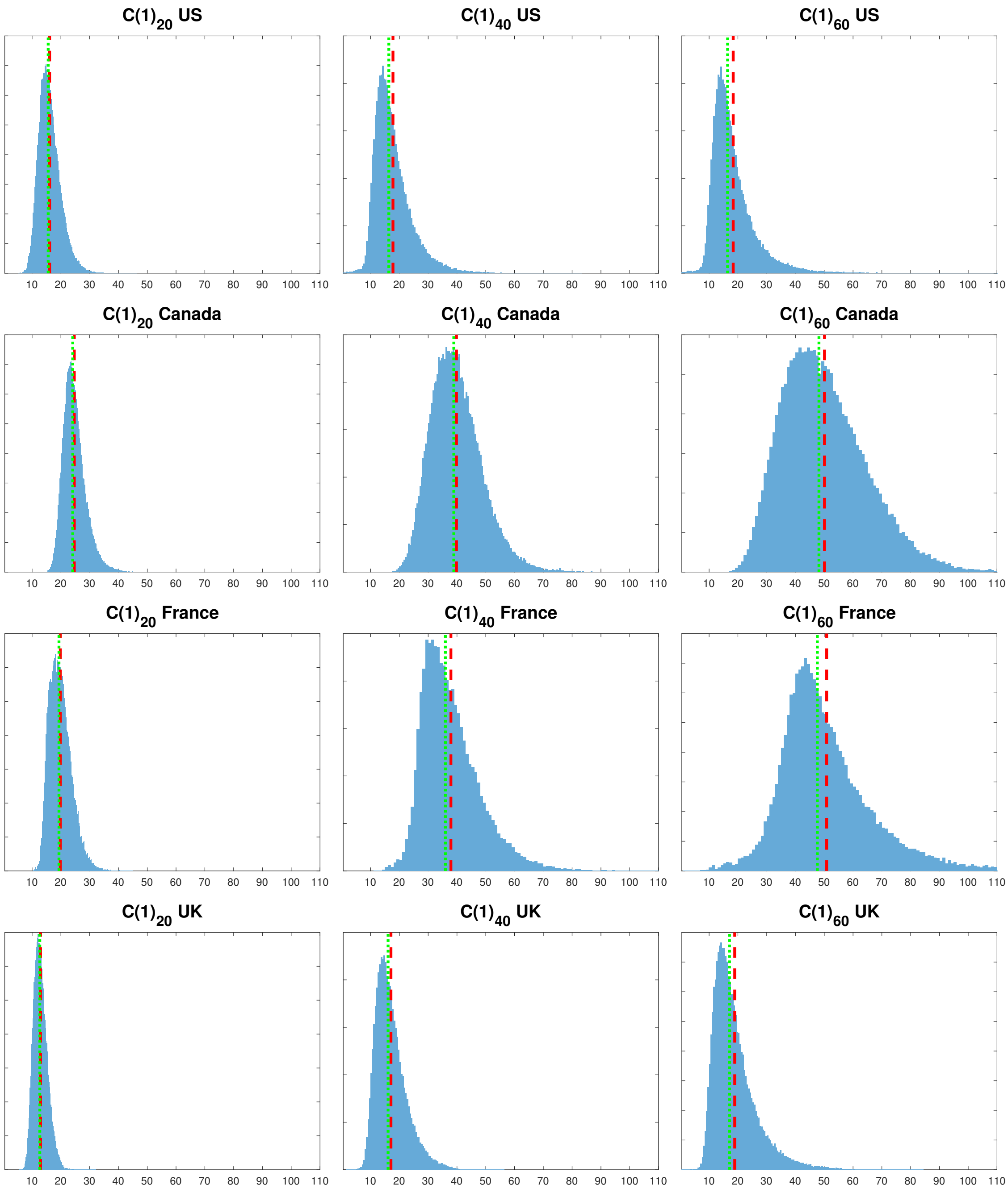

Figure 4.8: $C_n(1)$ for linear trend
Mean: Dashed line; Median: Dotted line

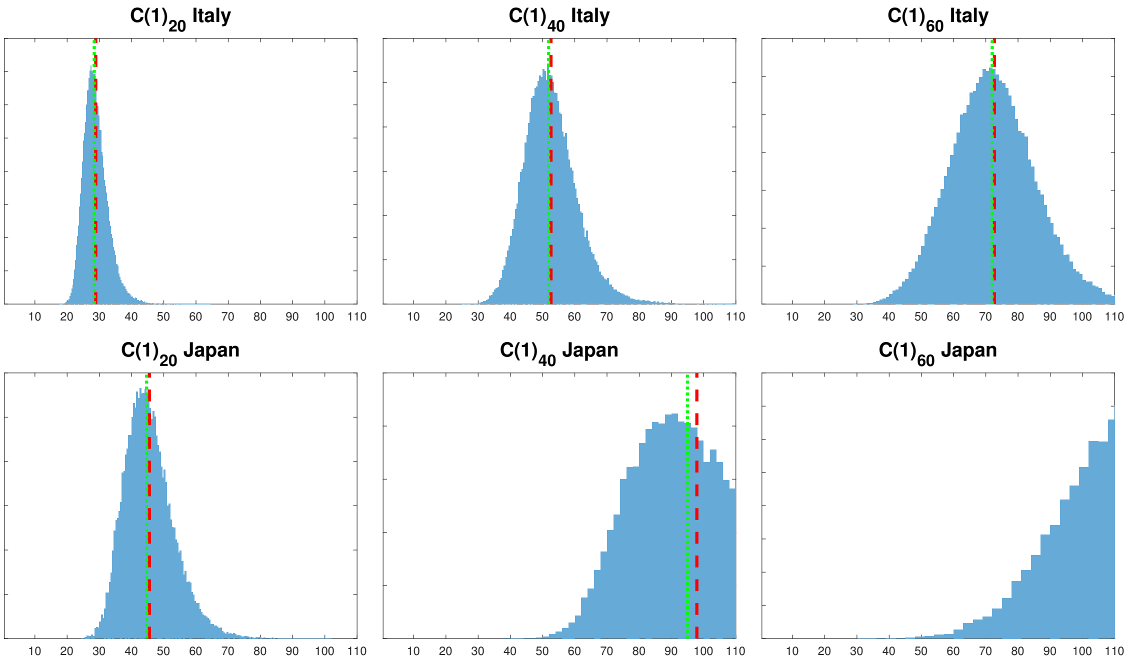

Figure 4.9: C(1)for linear trend
Mean: Dashed line; Median: Dotted line

Horizon	5	10	20	40	60
Canada	7.62 [7.55]	14 [13.7]	24.7 [24.1]	39.8 [38.9]	50.1 [48.2]
	[6.58; 8.9]	[11.5; 17.3]	[19.2; 31.9]	[27; 55.8]	[29.5; 76.9]
France	5 [4.96]	9.76 [9.65]	19.8 [19.3]	37.9 [36]	50.8 [47.6]
	[4.29; 5.81]	[7.86; 12.1]	[14.5; 26.8]	[25.4; 57.2]	[30.1; 82.4]
Italy	8.06 [8.02]	15.3 [15.1]	28.9 [28.4]	52.6 [51.9]	72.7 [72]
	[6.97; 9.28]	[12.8; 18.4]	[23.5; 35.8]	[40.5; 67.1]	[51.5; 96]
Japan	8.62 [8.6]	19.5 [19.4]	45.5 [44.7]	97.9 [95]	139 [133]
	[7.23; 10.1]	[15.3; 24.4]	[34.5; 59.5]	[69; 137]	[86.1; 212]
UK	5.34 [5.3]	8.64 [8.55]	12.9 [12.7]	17.1 [16.1]	18.9 [17.1]
	[4.77; 6.09]	[7.15; 10.4]	[9.39; 17.4]	[10.2; 27.5]	[10.3; 33.9]
US	7.6 [7.56]	12.4 [12.2]	16.1 [15.6]	17.8 [16.3]	18.4 [16.4]
	[6.55; 8.75]	[9.88; 15.4]	[10.7; 23.1]	[10.1; 30.4]	[10.2; 33.3]

Table 4.11: $C_n(1)$ for linear trend at different horizons; RJMCMC estimates
Mean, [median] with [90% credible sets] in the second row

Horizon	5	10	20	40	60
Canada	8.03; 7.27	14.6; 12.9	25.5; 22.4	40.4; 36.3	49.3; 45.3
France	4.4; 4.63	7.76; 9.15	13.6; 17.9	19.5; 28.4	15.7; 31.4
Italy	7.66; 7.66	13.9; 13.9	23.9; 23.9	32; 32	24.7; 24.7
Japan	7.45; 8.86	15.3; 19.6	32.8; 40.5	67.5; 69.4	90.6; 84.9
UK	6.11; 5.29	10.4; 8.58	13.6; 13.1	13.5; 17.3	13.5; 18.8
US	7.3; 7.3	11.3; 11.9	11.7; 13.1	9.51; 11.5	9.79; 11.7

Table 4.12: $C_n(1)$ for linear trend different horizons
Frequentist estimates for AIC; BIC

across countries than before. For example, at a horizon of 40 quarters the point estimates for the US are no longer contained in the credible sets of Canada, France, Italy or Japan and vice versa. The clustering of estimates is still present especially at longer horizons, but the clustering is different. Canada and France, and US and UK, now form two pairs for which the estimates are virtually identical at longer horizons. Italy and Japan exhibit higher persistence without the estimates converging as the horizon increases. The frequentist estimates are no longer as close to the ones from RJMCMC as before, but the majority is still contained in the credible sets.

Horizon	20				40				60			
Estimate	Mean	Median	AIC	BIC	Mean	Median	AIC	BIC	Mean	Median	AIC	BIC
Canada	3	3	2	3	3	3	2	2	4	3	2	2
France	4	4	4	4	4	4	4	4	3	4	4	3
Italy	2	2	3	2	2	2	3	3	2	2	3	4
Japan	1	1	1	1	1	1	1	1	1	1	1	1
UK	6	6	4	5	6	6	5	5	5	5	5	5
US	5	5	6	5	5	5	6	6	6	6	6	6

Table 4.13: Ranking by persistence for linear trend

Table 4.13 presents the persistence ranking for the linear detrending case. The persistence ranking remains mostly unchanged. Japan maintains a comfortable first place, followed by

Italy which is not far from the third and fourth place, Canada and France respectively, both of which exhibit similar persistence. Only the ranking for the US is changed substantially, having been ranked around third place in the difference stationary case it is now in fifth and sixth place depending on the estimate. This ranking for the US is more consistent with the one from the results of Campbell and Mankiw (1989). The rankings from the frequentist approach are very similar to those obtained with RJMCMC.

Kolmogorov-Smirnov Test for $C_n(1)$ Table 4.14 presents the test statistic for a horizon of 40 quarters. Additional results for different horizons can be found in the appendix. Again, the Kolmogorov-Smirnov test rejects the null hypothesis at the 1% level for all country pairs. However, in this case the US and the UK seem to have the most similar posterior, followed by the pair formed by Canada and France.

	Canada	France	Italy	Japan	UK	US
Canada	0	0.14051 (*)	0.57909 (*)	0.96998 (*)	0.89649 (*)	0.86332 (*)
France	0.14051 (*)	0	0.64029 (*)	0.96119 (*)	0.87076 (*)	0.83552 (*)
Italy	0.57909 (*)	0.64029 (*)	0	0.91528 (*)	0.98728 (*)	0.9708 (*)
Japan	0.96998 (*)	0.96119 (*)	0.91528 (*)	0	0.99974 (*)	0.99769 (*)
UK	0.89649 (*)	0.87076 (*)	0.98728 (*)	0.99974 (*)	0	0.03921 (*)
US	0.86332 (*)	0.83552 (*)	0.9708 (*)	0.99769 (*)	0.03921 (*)	0

Table 4.14: K-S test for $C(1)_{40}$ for linear trend

Conclusion In conclusion, the differences in persistence and the ordering of persistence between countries appear to mostly carry over to the linear detrending perspective, albeit with some changes in the ranking and clustering. The substantial persistence in the impulse response functions indicates that difference stationary models may be better suited to parsimoniously capture the very persistent dynamics of most of the series.

HP Detrended Data

I shall now turn to an analysis of the results obtained using deviations from an HP trend. As will become clear in the following, the results using HP detrended data seem to be dominated by filtering artifacts and do not seem particularly reliable in terms of capturing actual features of the data. Given this, the discussion of the results will be kept rather concise.

Model Choice Figure 4.10 shows the familiar posterior distributions over model indicators for the six countries. The models chosen here are of much higher order than those in the previous two cases, leading also to significantly more dispersed posteriors. This dispersion is to be expected, seeing as the likelihood is a function of autocorrelations and higher-order ARMA models can exhibit quite similar autocorrelation patterns even if the number of parameters differs. Put differently, near-cancellation of roots is more pronounced in higher-order ARMA models, a well known phenomenon (see e.g. Campbell and Mankiw (1987)).

The tendency of the algorithm to prefer higher-order ARMA models appears to be due to the application of the HP Filter which is known to introduce significant filtering artifacts at business cycle frequencies as documented by King and Rebelo (1993) and Cogley and Nason (1995a). Indeed, the impulse response functions shown below exhibit oscillations and periodicity very similar to the results of Cogley and Nason (1995a) who show that the HP filter can introduce periodicity in artificial data even if the underlying data generating process is completely aperiodic.

One can observe that the differences in the posterior distributions for the model indicators are not as striking as in the two foregoing cases. The posterior for Italy is now the least dispersed with a clear mode at the ARMA(3,2) model, followed by France with mode at the ARMA(3,3) model. The posteriors for the other countries show a clustering of the samples along the diagonal running from $(p, q) = (0, 0)$ to $(10, 10)$, again a sign of root cancellation.

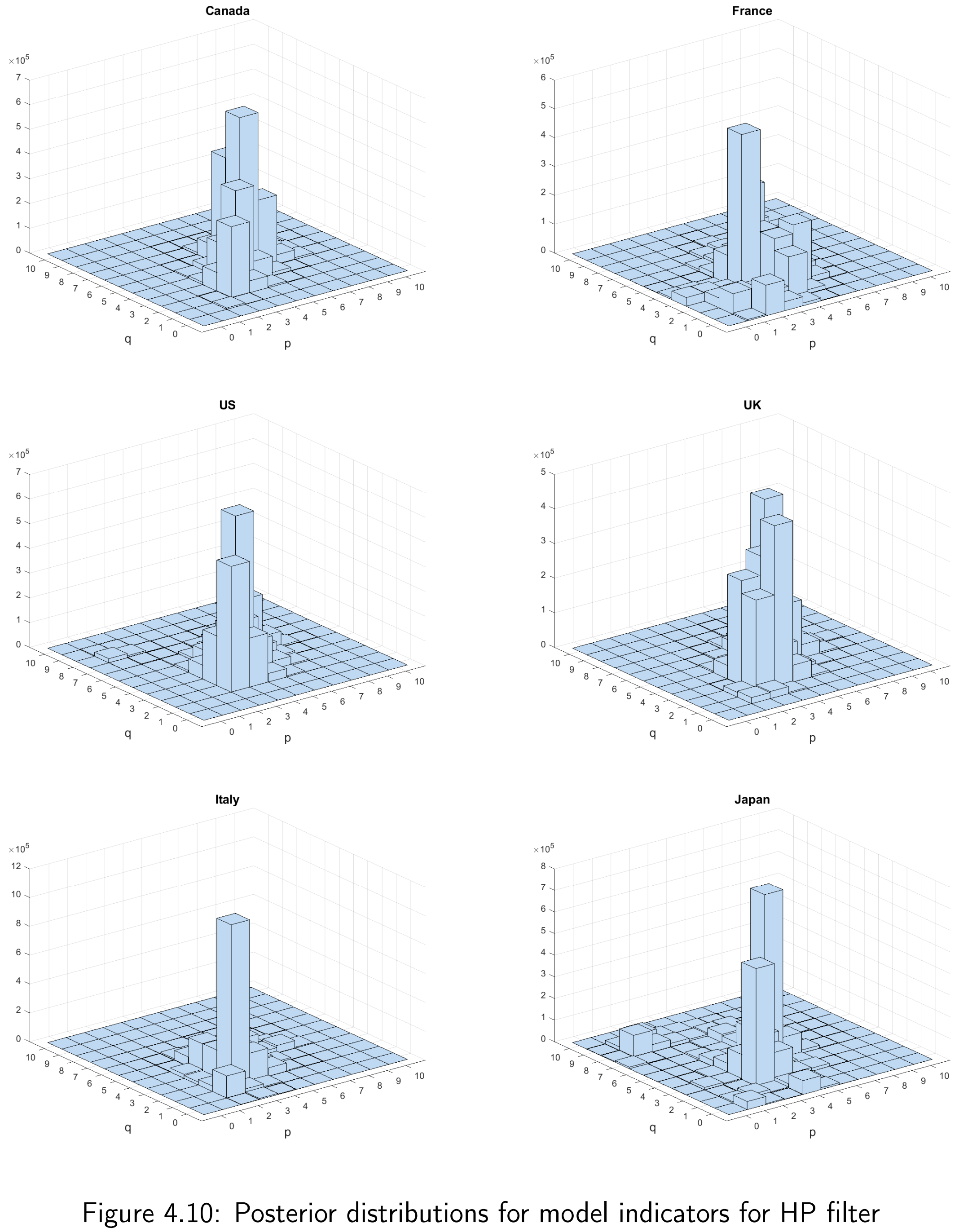

Figure 4.10: Posterior distributions for model indicators for HP filter

Impulse Responses Figure 4.11 presents the impulse responses. The frequentist estimates are shown in Table 4.15. Despite the substantial posterior uncertainty regarding the model choice, the credible sets for the impulse responses are surprisingly tight. Furthermore, the impulse response functions are quite similar across countries and exhibit clear cyclicality. The information criteria select models more in line with the results from RJMCMC. For all countries the responses are more or less identical to the mean and median response from RJMCMC while the response chosen by AIC and AICC for Japan are further away but still mostly contained in the credible set. Interestingly, the kink in the impulse response for France is still clearly visible, with the response dropping to about 20% of a shock standard deviation after 1 quarter.

All of the above suggests that the results from HP filtered data may indeed be an artifact of the filter chosen. Nevertheless, some insights may be obtained from analyzing the persistence measure as well as the corresponding ranking.

Persistence Figures 4.12 and 4.13 report the familiar posterior distributions for the persistence measure. Tables 4.16 and 4.17 report point estimates obtained from the posteriors and the frequentist estimates respectively.

The behavior of the mean and median estimates reflects the oscillations present in the impulse responses with the signs of the point estimates tending to change from positive to negative and back as the horizon increases. Notably, the posteriors for France and Japan exhibit a second mode at higher levels of persistence while all other posterior distributions of the persistence measure presented here and in the foregoing are unimodal.

Notably, while the means of the posteriors at a horizon of 20 quarters do not have the same sign, the medians are all negative. From the perspective of a zero-one loss function the cumulated response for all countries is thus first positive and then negative, only to turn positive or zero again.[6] The estimates, especially at longer horizons, are very much similar and all estimates are contained in the credible sets for all other countries starting at a horizon

[6] This difference in the means and medians is a result of multi-modality and skewness in the posterior distributions.

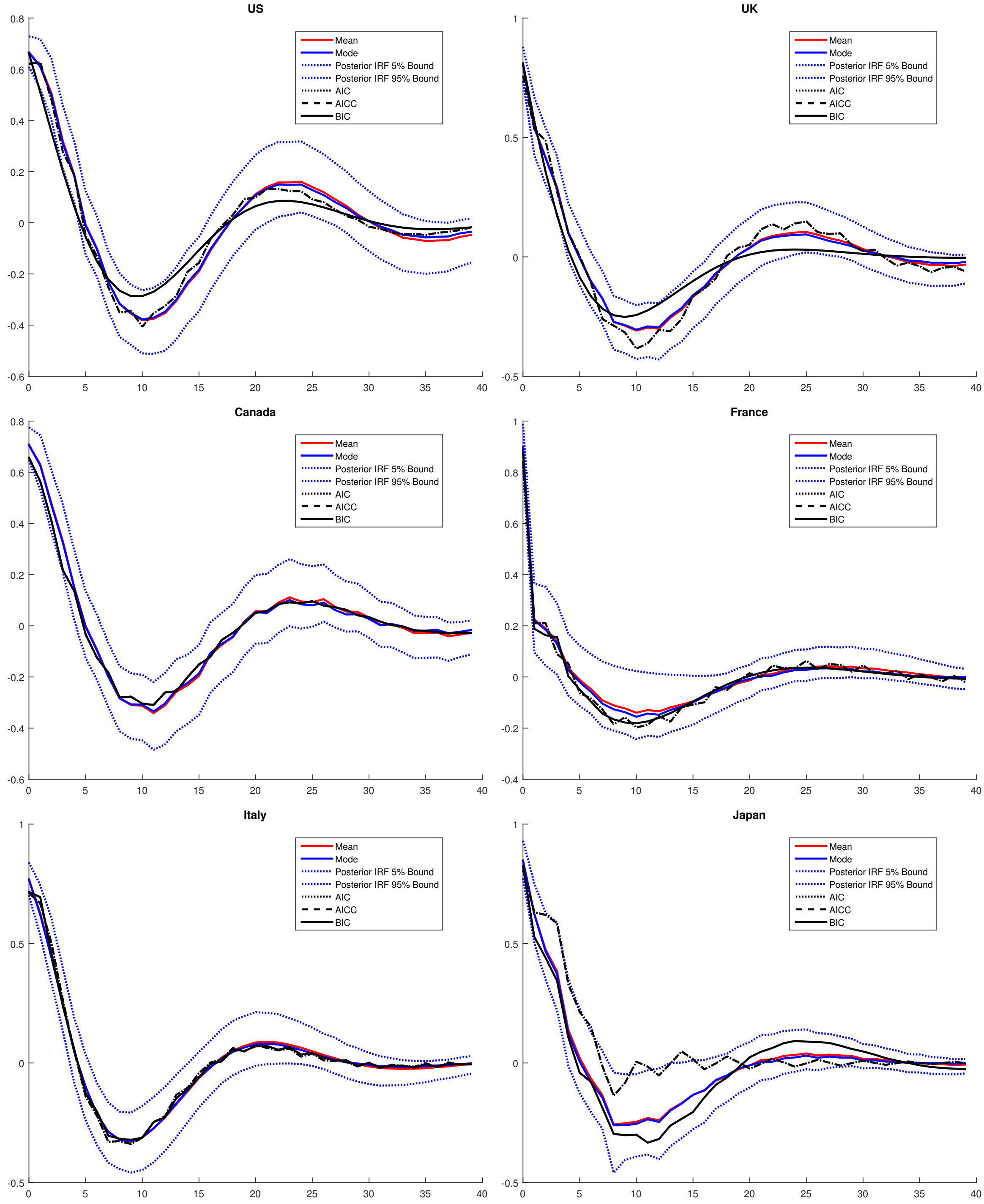

Figure 4.11: Estimated impulse responses for HP filter

Country	Criterion	P_1	P_2	P_3	P_4	P_5	Q_1	Q_2	Q_3	Q_4	Q_5	σ_e
Canada	AIC	0.423	0.744	0.377	−0.719		0.435	−0.487	−0.948			0.657
		(0.032)	(0.037)	(0.025)	(0.019)		(0.045)	(0.056)	(0.043)			(0.044)
	AICC	0.423	0.744	0.377	−0.719		0.435	−0.487	−0.948			0.657
		(0.032)	(0.037)	(0.025)	(0.019)		(0.045)	(0.056)	(0.043)			(0.044)
	BIC	0.423	0.744	0.377	−0.719		0.435	−0.487	−0.948			0.657
		(0.032)	(0.037)	(0.025)	(0.019)		(0.045)	(0.056)	(0.043)			(0.044)
France	AIC	0.141	0.847	0.694	−0.694	−0.173	0.108	−0.633	−0.833	0.358		0.845
		(0.142)	(0.054)	(0.077)	(0.093)	(0.084)	(0.177)	(0.107)	(0.090)	(0.160)		(0.068)
	AICC	0.141	0.847	0.694	−0.694	−0.173	0.108	−0.633	−0.833	0.358		0.845
		(0.142)	(0.054)	(0.077)	(0.093)	(0.084)	(0.177)	(0.107)	(0.090)	(0.160)		(0.068)
	BIC	1.215	−0.075	−0.032	−0.192		−1.000					0.877
		(0.055)	(0.052)	(0.111)	(0.075)		(0.051)					(0.035)
Italy	AIC	0.003	1.116	0.320	−0.718		0.934	−0.417	−1.000	−0.298	−0.159	0.712
		(0.073)	(0.051)	(0.056)	(0.065)		(0.076)	(0.115)	(0.077)	(0.109)	(0.093)	(0.048)
	AICC	0.003	1.116	0.320	−0.718		0.934	−0.417	−1.000	−0.298	−0.159	0.712
		(0.073)	(0.051)	(0.056)	(0.065)		(0.076)	(0.115)	(0.077)	(0.109)	(0.093)	(0.048)
	BIC	0.224	1.123	0.086	−0.791	0.145	0.745	−0.706	−0.986			0.716
		(0.055)	(0.035)	(0.063)	(0.038)	(0.050)	(0.037)	(0.030)	(0.032)			(0.048)
Japan	AIC	1.195	−1.317	1.000	−0.413		−0.414	1.150	−0.163	0.183	0.287	0.808
		(0.218)	(0.193)	(0.194)	(0.153)		(0.210)	(0.090)	(0.252)	(0.089)	(0.088)	(0.068)
	AICC	1.195	−1.317	1.000	−0.413		−0.414	1.150	−0.163	0.183	0.287	0.808
		(0.218)	(0.193)	(0.194)	(0.153)		(0.210)	(0.090)	(0.252)	(0.089)	(0.088)	(0.068)
	BIC	1.371	−0.846	1.000	−0.634		−0.730	0.501	−0.771			0.824
		(0.117)	(0.233)	(0.178)	(0.064)		(0.104)	(0.173)	(0.111)			(0.062)
UK	AIC	0.409	0.231	0.758	−0.102	−0.546	0.295	0.121	−0.823	−0.593		0.759
		(0.145)	(0.138)	(0.042)	(0.127)	(0.112)	(0.173)	(0.057)	(0.042)	(0.149)		(0.041)
	AICC	0.409	0.231	0.758	−0.102	−0.546	0.295	0.121	−0.823	−0.593		0.759
		(0.145)	(0.138)	(0.042)	(0.127)	(0.112)	(0.173)	(0.057)	(0.042)	(0.149)		(0.041)
	BIC	1.702	−0.757				−1.000					0.810
		(0.021)	(0.013)				(0.037)					(0.038)
US	AIC	0.270	1.017	0.232	−0.741		0.698	−0.486	−0.989	−0.131	−0.092	0.622
		(0.037)	(0.054)	(0.026)	(0.018)		(0.082)	(0.104)	(0.071)	(0.088)	(0.079)	(0.037)
	AICC	0.270	1.017	0.232	−0.741		0.698	−0.486	−0.989	−0.131	−0.092	0.622
		(0.037)	(0.054)	(0.026)	(0.018)		(0.082)	(0.104)	(0.071)	(0.088)	(0.079)	(0.037)
	BIC	1.770	−0.831				−1.000					0.661
		(0.015)	(0.008)				(0.024)					(0.035)

Table 4.15: Frequentist regression results for HP filter

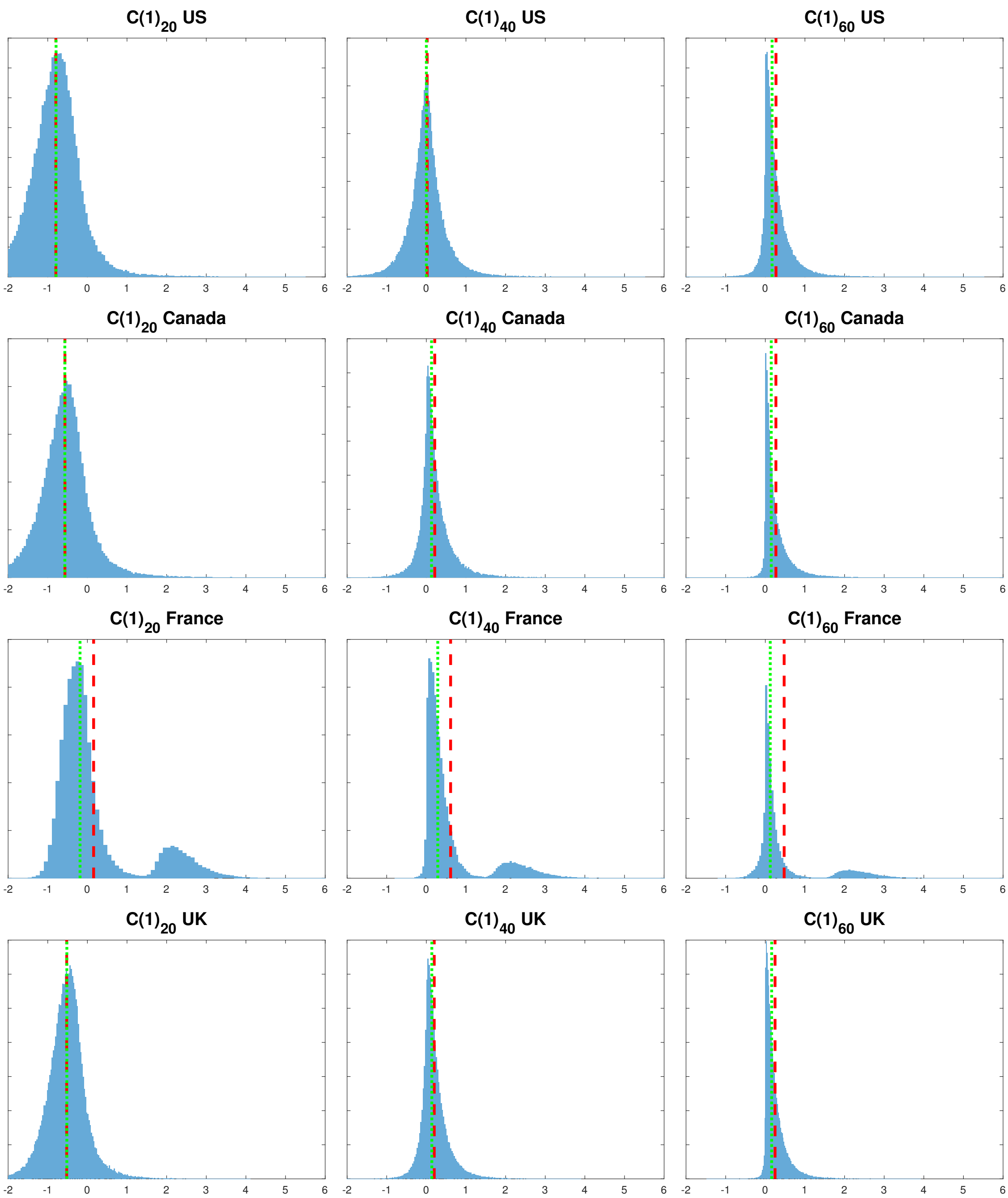

Figure 4.12: $C_n(1)$ for HP filter
Mean: Dashed line; Median: Dotted line

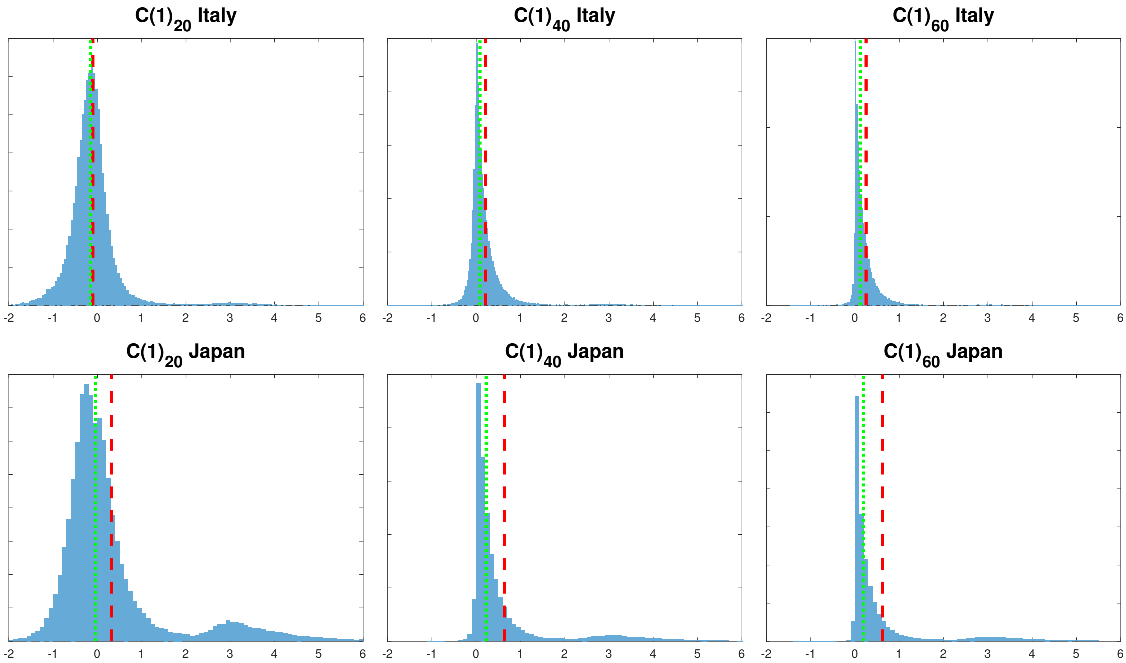

Figure 4.13: $C_n(1)$ for HP filter
Mean: Dashed line; Median: Dotted line

of 10 quarters. At a horizon of 60 quarters, the point estimates for Canada, Italy, the UK, and the US are virtually identical.

The frequentist estimates for the persistence are mostly in line with expectations formed during inspection of the impulse responses and AIC and BIC tend to deliver similar estimates with the exception of Japan. While for Japan the impulse response functions already hint at significantly different persistence estimates from AIC and BIC, the difference in the impulse response functions chosen by the different criteria is not as pronounced for France. Nonetheless, the point estimates differ significantly for the latter country with the AIC estimate at a horizon of 60 quarters being 3.82 while the model chosen by the BIC implies an estimate of 0.0142. In general, the persistence estimates at a horizon of 60 quarters are zero for almost all countries and criteria while they are clearly positive for RJMCMC.

Despite the aforementioned considerations challenging the dependability of the results,

Horizon	5	10	20	40	60
Canada	3.24 $[3.21]$	1.55 $[1.46]$	-0.567 $[-0.569]$	0.212 $[0.131]$	0.268 $[0.15]$
	$[2.61; 3.99]$	$[0.554; 2.85]$	$[-1.6; 0.459]$	$[-0.316; 0.977]$	$[-0.00424; 0.944]$
France	1.63 $[1.6]$	1.06 $[0.87]$	0.157 $[-0.185]$	0.613 $[0.289]$	0.476 $[0.125]$
	$[1.1; 2.27]$	$[0.3; 2.49]$	$[-0.804; 2.54]$	$[0.019; 2.55]$	$[-0.157; 2.55]$
Italy	2.65 $[2.62]$	0.739 $[0.661]$	-0.102 $[-0.153]$	0.208 $[0.0852]$	0.25 $[0.119]$
	$[2.05; 3.36]$	$[-0.159; 1.88]$	$[-0.856; 0.648]$	$[-0.166; 0.847]$	$[-0.0102; 0.83]$
Japan	2.93 $[2.87]$	1.8 $[1.6]$	0.316 $[-0.0461]$	0.642 $[0.225]$	0.618 $[0.188]$
	$[2.31; 3.78]$	$[0.751; 3.63]$	$[-0.873; 3.41]$	$[-0.00059; 3.4]$	$[0.00408; 3.4]$
UK	2.68 $[2.65]$	1.25 $[1.19]$	-0.522 $[-0.518]$	0.199 $[0.139]$	0.246 $[0.161]$
	$[2.17; 3.27]$	$[0.467; 2.26]$	$[-1.3; 0.243]$	$[-0.232; 0.816]$	$[0.00434; 0.788]$
US	3.43 $[3.39]$	1.34 $[1.26]$	-0.796 $[-0.79]$	0.0211 $[0.00077]$	0.27 $[0.171]$
	$[2.8; 4.18]$	$[0.367; 2.61]$	$[-1.85; 0.217]$	$[-0.73; 0.83]$	$[-0.0994; 0.961]$

Table 4.16: $C_n(1)$ for HP filter at different horizons; RJMCMC estimates
Mean, [median] with [90% credible sets] in the second row

Horizon	5	10	20	40	60
Canada	$2.96;\ 2.96$	$1.19;\ 1.19$	$-0.827;\ -0.827$	$0.0119;\ 0.0119$	$0.026;\ 0.026$
France	$1.59;\ 1.53$	$0.703;\ 0.653$	$-0.398;\ -0.316$	$0.0548;\ 0.034$	$-0.0009;\ -0.0001$
Italy	$2.88;\ 2.84$	$0.728;\ 0.799$	$-0.0472;\ -0.0596$	$0.199;\ 0.231$	$0.211;\ 0.25$
Japan	$3.95;\ 2.67$	$3.85;\ 1.26$	$3.81;\ -0.743$	$3.82;\ 0.0648$	$3.82;\ 0.0142$
UK	$2.84;\ 2.28$	$1.05;\ 0.898$	$-0.967;\ -0.281$	$-0.0512;\ 0.014$	$0.0583;\ 0.0002$
US	$3.45;\ 2.62$	$1.06;\ 0.786$	$-0.827;\ -0.618$	$-0.141;\ -0.0441$	$0.0035;\ 0.0137$

Table 4.17: $C_n(1)$ for HP filter at different horizons
Frequentist estimates for AIC; BIC

Horizon	20				40				60			
Estimate	Mean	Median	AIC	BIC	Mean	Median	AIC	BIC	Mean	Median	AIC	BIC
Canada	5	5	4	6	3	4	4	5	4	4	4	2
France	2	3	3	3	2	1	3	3	2	5	6	6
Italy	3	2	2	1	4	5	2	1	5	6	2	1
Japan	1	1	1	5	1	2	1	2	1	1	1	3
UK	4	4	6	2	5	3	5	4	6	3	3	5
US	6	6	4	4	6	6	6	6	3	2	5	4

Table 4.18: Ranking by persistence for HP filter

Table 4.18 reports the same ranking as in the foregoing.[7] Due to the multimodal nature of some of the posteriors, the rankings do not coincide across mean and median based estimates as they do in the previous sections. Furthermore, especially at longer horizons, the estimates are almost identical with each of the estimates captured in the credible sets of all the others. France appears somewhat more persistent as before and the US experiences an "improvement" in its persistence ranking as the horizon increases, moving from sixth to third (second) place in the ranking of the means (medians). The rankings do not coincide between the different methods as well as before.

Kolmogorov Smirnov Test for $C_n(1)$ Again, the Kolmogorov-Smirnov test rejects the null hypothesis of equality of the posterior distributions for all country pairs and horizons at the one percent level. Table 4.19 reports the test statistic for a horizon of 40 quarters. Additional tables for different horizons can be found in the appendix. In the HP detrended case, the closest two distributions are now those for Canada and the US, with the pairs Italy and Canada and Canada and Japan following in terms of magnitude of the test statistic.

[7] It is not clear how the negative estimates are to be treated in this context. What does a negative estimate tell us? Is a negative estimate more or less persistent than a positive estimate of the same magnitude? The ranking presented here just reflects the arrangement of the estimates on the real line.

	Canada	France	Italy	Japan	UK	US
Canada	0	0.25295 (*)	0.08726 (*)	0.19597 (*)	0.02858 (*)	0.25459 (*)
France	0.25295 (*)	0	0.33926 (*)	0.1054 (*)	0.23759 (*)	0.47643 (*)
Italy	0.08726 (*)	0.33926 (*)	0	0.24071 (*)	0.10391 (*)	0.28712 (*)
Japan	0.19597 (*)	0.1054 (*)	0.24071 (*)	0	0.17015 (*)	0.44848 (*)
UK	0.02858 (*)	0.23759 (*)	0.10391 (*)	0.17015 (*)	0	0.28146 (*)
US	0.25459 (*)	0.47643 (*)	0.28712 (*)	0.44848 (*)	0.28146 (*)	0

Table 4.19: K-S test for $C(1)_{40}$ for HP filter

Conclusion To conclude, the validity of the results using HP filtered data is uncertain. The impulse responses show a cyclicality which may very well be introduced by the filter, making any estimate of persistence, at the very least, less reliable. The rankings between countries appear less consistent compared to the previous sections and the posterior distributions exhibit multi-modality making the choice between means and medians more onerous. Nonetheless, even when applying a filter that is designed to filter out low-frequency movement in the data, some persistence remains even at long horizons in the RJMCMC estimates and the behavior of the economies differs, albeit not as strongly as before.

4.12 US GDP Components

In this section, the dynamics of the major components of GDP– private consumption, gross fixed capital formation, government consumption, imports, and exports– in the US are analyzed in isolation in order to gain insight into which of the components are the main drivers behind the above results.

The data used in this section is again the VOBARSA measure, that is, seasonally adjusted volume estimates, taken from the OECD.stat website for the period 1960:1 to 2007:4. The data was transformed into per capita terms and first differences of the logarithms were taken as in the foregoing. The sampler settings were adjusted for each series, again using short pilot

Component	Object	Proposal
Capital Formation	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.08^2)$
	σ_ϵ	$TN(\mu, 0.05^2)$
Exports	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.12^2)$
	σ_ϵ	$TN(\mu, 0.05^2)$
Government Consumption	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.06^2)$
	σ_ϵ	$TN(\mu, 0.05^2)$
Imports	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.055^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.1^2)$
	σ_ϵ	$TN(\mu, 0.07^2)$
Private Consumption	p	$DL(\mu, 2.2)$
	q	$DL(\mu, 2.2)$
	(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
	(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.06^2)$
	σ_ϵ	$TN(\mu, 0.05^2)$

Table 4.20: Proposal distributions for GDP components

	α	α_w	α_b
Exports	0.28	0.37	0.09
Government Consumption	0.27	0.36	0.08
Gross Fixed Capital Formation	0.20	0.27	0.06
Imports	0.64	0.79	0.13
Private Consumption	0.26	0.35	0.08

Table 4.21: Acceptance rates for GDP components
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

runs. The chosen parameter values are presented in Table 4.20[8] and the resulting acceptance rates are contained in Table 4.21.

4.12.1 Model Choice

The posteriors for the model indicators presented in Figure 4.14 show quite intriguing differences. The posterior for imports is the least dispersed with a clear mode at the random walk. The mode model for the exports series is clearly AR(1) but there is some posterior uncertainty around that point. The mode for the government consumption series is at the ARMA(1,1) model with a medium level of posterior uncertainty. The other two posteriors for private consumption and gross fixed capital formation show substantially higher posterior model uncertainty. The quite pronounced mode for the capital formation series is at the MA(2) model, but the posterior is very dispersed with samples even for high-order models like ARMA(6,4). The posterior distribution for private consumption is not quite as dispersed and does not exhibit as clear a mode as the one for capital formation. The mode for this series lies at the ARMA(1,1) model but e.g. the AR(2) model is attached an only slightly lower posterior probability.

[8]The notation here is the same as in table 4.2.

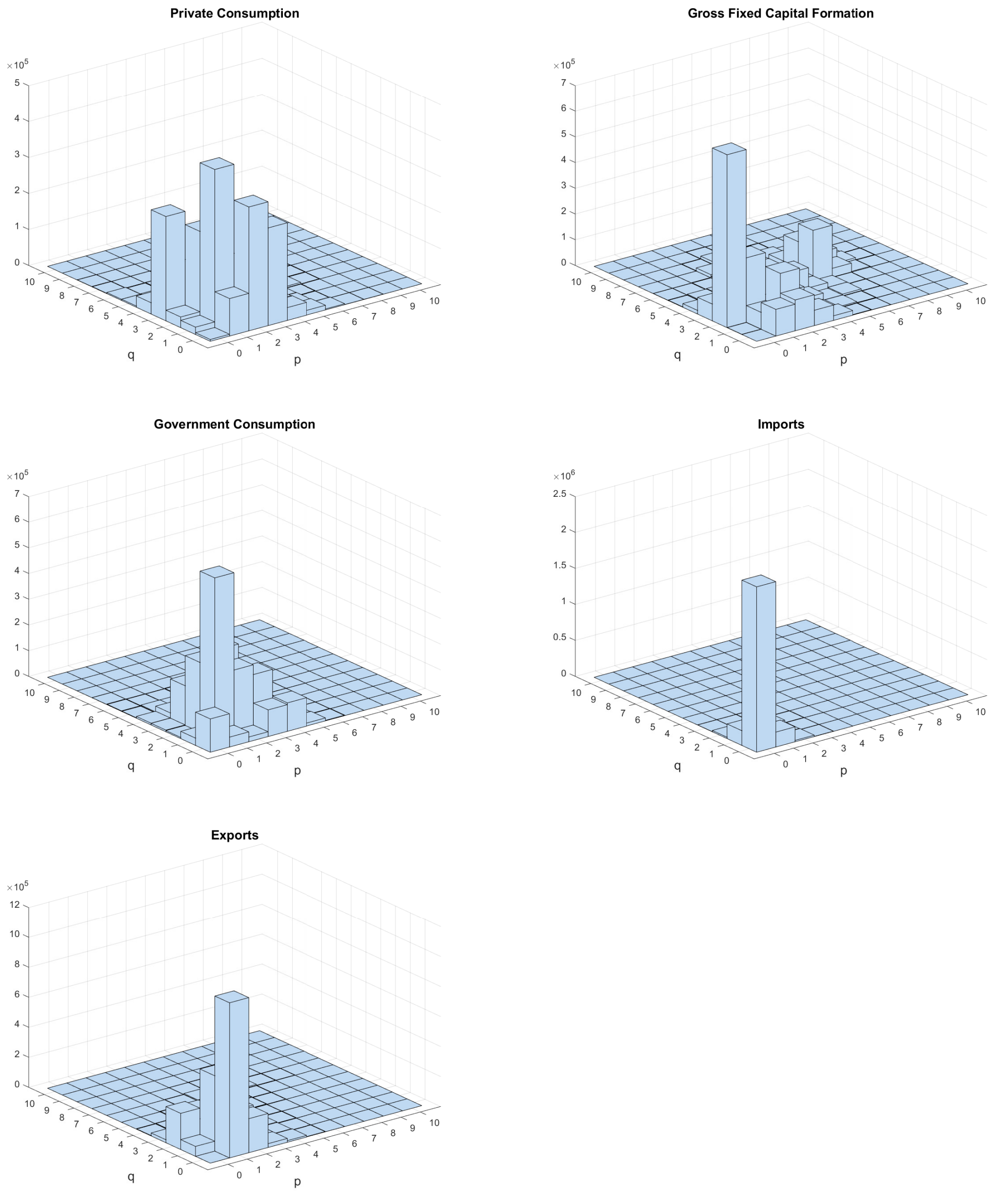

Figure 4.14: Posterior distributions for model indicators for GDP components

4.12.2 Impulse Responses

The impulse responses of the GDP components are presented in Figure 4.15. Table 4.22 contains the results from the frequentist regressions. Credible sets for the impulse responses are tight, with somewhat more uncertainty in the estimates for capital formation and government consumption. The clearly preferred model for the imports series is a pure random walk for all methods, which is reflected in the shape of the impulse response. The credible sets contain responses from some samples with AR and MA models of order one respectively. The posterior for the exports series exhibits the exponential decay from the AR(1) model at the mode with oscillatory behavior. All three information criteria pick a model with oscillatory behavior for this series, while only the AIC and AICC estimates show high frequency oscillations for private consumption and government consumption and a low frequency cycle for capital formation. The persistence in the growth rate for the exports series is relatively limited based on the impulse response, but exports as well as imports have by far the greatest shock standard deviation at about 3.5 percentage points, followed by capital formation with about 1.5 percentage points.

The impulse responses for the two consumption and the capital formation series show a somewhat more intricate behavior. Private consumption exhibits medium persistence. The oscillations from the model picked by the AIC and AICC are present in the RJMCMC estimates only to a very limited extent in the shape of the credible sets. Rather, the impulse responses from RJMCMC and BIC decay exponentially after the effect of the low-order MA terms vanishes. The impulse response for government consumption follows a similar pattern, although the impulse responses from AIC and AICC, which again show oscillatory behavior, are more persistent than the ones chosen by either RJMCMC or BIC.

The impulse response for capital formation reflects the shape of the posterior over the model orders in the shape and width of the credible sets and the behavior of the mean and median responses. While the median response is zero after 5 quarters, the mean response stays negative until quarter 20 after the shock. The BIC chooses the rather simple MA(2) model, and

Component	Criterion	P_1	P_2	P_3	P_4	P_5	Q_1	Q_2	Q_3	Q_4	Q_5	σ_e
Exports	AIC	−0.043	0.909				−0.211	−0.871	0.319	−0.128	−0.106	3.220
		(0.048)	(0.047)				(0.070)	(0.085)	(0.094)	(0.075)	(0.087)	(0.792)
	AICC	−0.043	0.909				−0.211	−0.871	0.319	−0.128	−0.106	3.220
		(0.048)	(0.047)				(0.070)	(0.085)	(0.094)	(0.075)	(0.087)	(0.792)
	BIC	−0.973					0.751	−0.139	0.191	0.080		3.263
		(0.003)					(0.047)	(0.080)	(0.092)	(0.081)		(0.830)
Government Consumption	AIC	0.920	−0.358	0.382	−0.862	0.699	−0.769	0.277	−0.211	0.983	−0.790	0.789
		(0.083)	(0.037)	(0.042)	(0.036)	(0.069)	(0.109)	(0.051)	(0.056)	(0.043)	(0.111)	(0.064)
	AICC	0.920	−0.358	0.382	−0.862	0.699	−0.769	0.277	−0.211	0.983	−0.790	0.789
		(0.083)	(0.037)	(0.042)	(0.036)	(0.069)	(0.109)	(0.051)	(0.056)	(0.043)	(0.111)	(0.064)
	BIC	0.920	−0.358	0.382	−0.862	0.699	−0.769	0.277	−0.211	0.983	−0.790	0.789
		(0.083)	(0.037)	(0.042)	(0.036)	(0.069)	(0.109)	(0.051)	(0.056)	(0.043)	(0.111)	(0.064)
Gross Fixed Capital Formation	AIC	1.808	−0.846				−1.518	0.543	−0.242	0.380	−0.163	1.683
		(0.022)	(0.015)				(0.077)	(0.128)	(0.123)	(0.125)	(0.088)	(0.227)
	AICC	1.808	−0.846				−1.518	0.543	−0.242	0.380	−0.163	1.683
		(0.022)	(0.015)				(0.077)	(0.128)	(0.123)	(0.125)	(0.088)	(0.227)
	BIC						0.371	0.308				1.763
							(0.068)	(0.060)				(0.216)
Imports	AIC											3.412
												(0.668)
	AICC											3.412
												(0.668)
	BIC											3.412
												(0.668)
Private Consumption	AIC	−0.479	−0.453	0.432	0.216		0.676	0.825				0.622
		(0.120)	(0.074)	(0.059)	(0.081)		(0.102)	(0.061)				(0.034)
	AICC	−0.479	−0.453	0.432	0.216		0.676	0.825				0.622
		(0.120)	(0.074)	(0.059)	(0.081)		(0.102)	(0.061)				(0.034)
	BIC	0.191	0.196									0.648
		(0.068)	(0.064)									(0.033)

Table 4.22: Frequentist regression results for GDP components

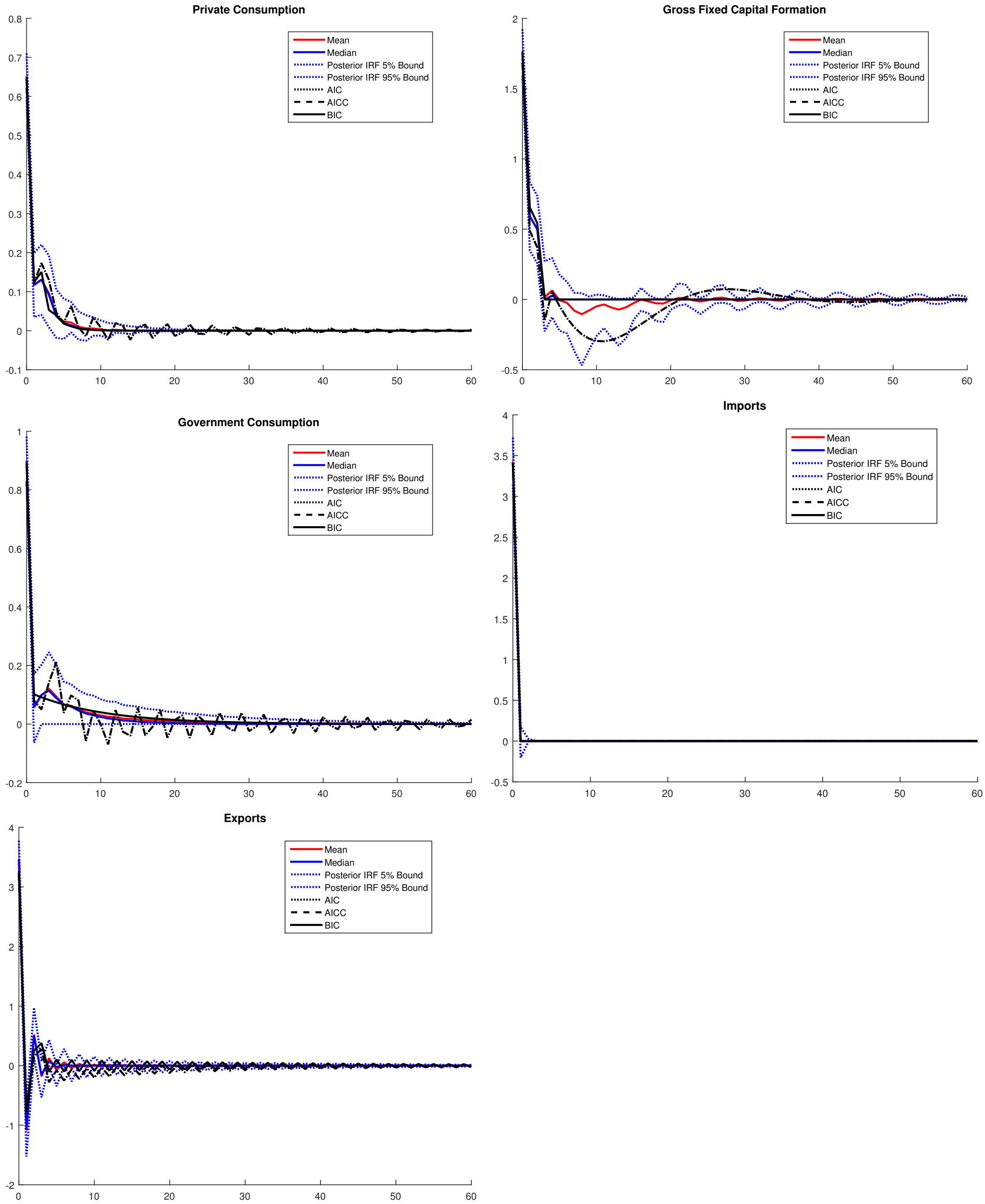

Figure 4.15: Estimated impulse responses for GDP components

the models chosen by the AIC and AICC show persistent oscillation. This oscillatory behavior is present to some degree in the mean response from RJMCMC as well as the credible sets.

Judging from the perspective of impulse responses alone, the shape of the impulse response for the two consumption series is closest to the one for the whole economy. This is not entirely surprising as these two components account for a significant proportion of GDP. They cannot, however, account for the negative responses contained in the credible set for the full GDP series. This feature could, however, be explained by the negative response of capital formation.

4.12.3 Persistence

Turning to the analysis of the posteriors for the persistence measure for the series, presented in figures 4.16. Table 4.23 presents point estimates of $C_n(1)$ from RJMCMC. The frequentist estimates can be found in Table 4.24.

The plot of $C_n(1)$ for the capital formation series immediately stands out. The distribution for capital formation is significantly bi-modal and very dispersed with substantial probability mass at $C_n(1) = 0$, possibly indicating some degree of trend reversion. The shape of the posterior is reflected in the point estimates, both from RJMCMC and the information criteria. The estimate from the AIC is almost zero at a horizon of 60 quarters while the one from BIC equals 1.68, roughly in line with the median estimate from RJMCMC, which is equal to 1.58.

Similarly differing estimates are obtained for the exports series, with the AIC estimate at 0.036 and the BIC estimate at 0.95. For imports, all methods agree on the pure random walk model resulting in persistence estimates equal to one.

The posterior for government consumption exhibits a peak at $C_n(1) = 1$, a consequence of the presence of some pure random walk models in the posterior. The point estimates for government consumption indicate substantial persistence, however, with a mean estimate from RJMCMC of 2.03 at a horizon of 60 quarters. The frequentist estimates are similar for this series. Private consumption is not quite as persistent with a mean of 1.7 at the same horizon with the frequentist estimates bracketing this value at 1.95 and 1.63 for AIC and BIC

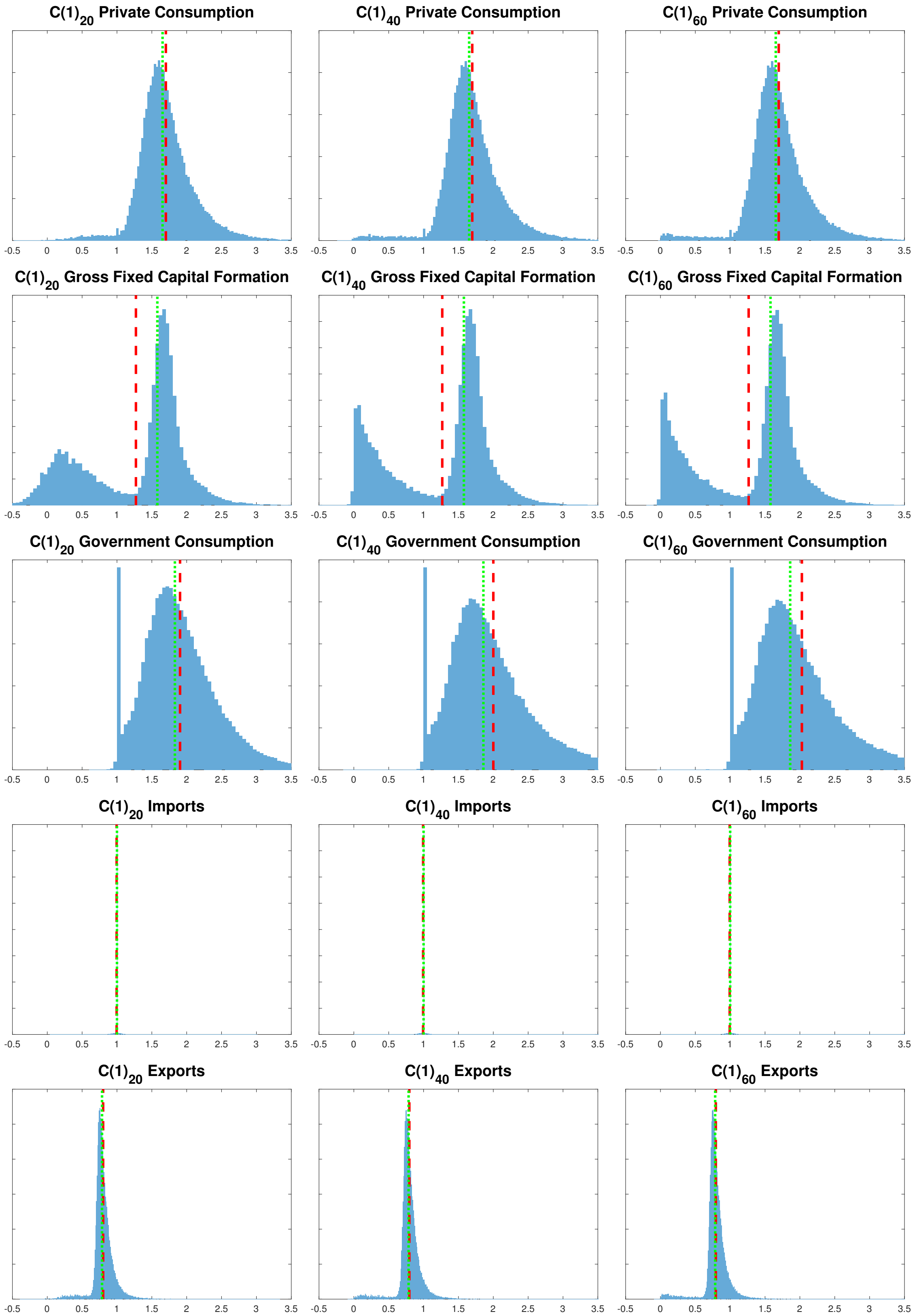

Figure 4.16: $C_n(1)$ for GDP components
Mean: Dashed line; Median: Dotted line

respectively. Again, the estimates using BIC are closest to those obtained with RJMCMC.

Horizon	5	10	20	40	60
Exports	0.8 $[0.78]$	0.809 $[0.787]$	0.801 $[0.786]$	0.796 $[0.785]$	0.794 $[0.785]$
	$[0.676; 0.993]$	$[0.674; 1.03]$	$[0.649; 1.04]$	$[0.621; 1.04]$	$[0.603; 1.04]$
Government	1.49 $[1.49]$	1.73 $[1.71]$	1.9 $[1.83]$	2 $[1.86]$	2.03 $[1.86]$
Consumption	$[1.05; 1.85]$	$[1.05; 2.37]$	$[1.04; 2.92]$	$[1.04; 3.33]$	$[1.04; 3.45]$
Capital Formation	1.66 $[1.66]$	1.46 $[1.59]$	1.27 $[1.58]$	1.27 $[1.58]$	1.26 $[1.58]$
	$[1.25; 2.1]$	$[0.582; 2.16]$	$[-0.0151; 2.17]$	$[0.0587; 2.17]$	$[0.0574; 2.17]$
Imports	0.998 $[1]$	0.997 $[1]$	0.995 $[1]$	0.993 $[1]$	0.993 $[1]$
	$[0.931; 1.06]$	$[0.928; 1.06]$	$[0.928; 1.06]$	$[0.928; 1.06]$	$[0.927; 1.06]$
Private	1.62 $[1.61]$	1.69 $[1.65]$	1.7 $[1.65]$	1.7 $[1.65]$	1.7 $[1.65]$
Consumption	$[1.27; 2.03]$	$[1.23; 2.31]$	$[1.18; 2.42]$	$[1.17; 2.44]$	$[1.17; 2.44]$

Table 4.23: $C_n(1)$ for GDP components at different horizons; RJMCMC estimates
Mean, [median] with [90% credible sets] in the second row

Horizon	5	10	20	40	60
Exports	$0.769;\ 0.97$	$0.525;\ 0.941$	$0.269;\ 0.944$	$0.0805;\ 0.948$	$0.0358;\ 0.951$
Government Consumption	$1.63;\ 1.47$	$1.83;\ 1.75$	$1.75;\ 2.01$	$1.77;\ 2.14$	$1.77;\ 2.16$
Capital Formation	$1.43;\ 1.68$	$0.741;\ 1.68$	$-0.351;\ 1.68$	$0.0791;\ 1.68$	$-0.0123;\ 1.68$
Imports	$1;\ 1$	$1;\ 1$	$1;\ 1$	$1;\ 1$	$1;\ 1$
Private Consumption	$1.79;\ 1.59$	$1.95;\ 1.63$	$1.94;\ 1.63$	$1.94;\ 1.63$	$1.95;\ 1.63$

Table 4.24: $C_n(1)$ for GDP components at different horizons
Frequentist estimates for AIC; BIC

4.12.4 Summary

In conclusion, the persistence and shape of the impulse response of the GDP series seems to be driven mainly by the two consumption series. Regarding the inclusion of negative responses in the credible sets for the aggregate series, it can be conjectured that this phenomenon may be explained by the response of capital formation since the shape and persistence of the line traced out by the lower 5% credible set bound is reminiscent of the shape of the response in the latter series. Furthermore, none of the other substantial series show meaningful negative responses, neither with respect to magnitude nor posterior mass.

4.13 UK Subsamples

The result for the UK GDP series appears quite curious. The clear preference for a pure random walk may indicate that the likelihood is dominated by rare and substantial shifts in the level of GDP which are not well captured by adding persistence through the growth rate. In order to gain some insight into the validity of this conjecture, the series for the UK was divided into two subsamples at two different points in time. The first break point chosen is the beginning of the year 1980, corresponding roughly to the assumption of office by Margaret Thatcher. The second break point chosen is the fourth quarter of 1989 corresponding to the end of Margaret Thatcher's time in office as well as the collapse of the Soviet Union.

Sampler settings were the same as for all other estimations for first differences. The resulting acceptance rates are presented in Table 4.25.

4.13.1 Model Choice

Figure 4.17 presents the posterior distributions of the model indicators for the subsamples. While the posterior for the subsample stretching from 1960:1 to 1989:4 strongly resembles the one for the whole series with clear preference for a pure random walk and only a few samples with low-order AR and MA models, the posterior for the subsample for the period 1960:1

	α	α_w	α_b
1960:1 - 1979:4	0.48	0.63	0.13
1980:1 - 2007:4	0.29	0.39	0.09
1960:1 - 1989:1	0.55	0.68	0.15
1990:1 - 2007:4	0.30	0.39	0.09

Table 4.25: Acceptance rates for UK subsamples
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

to 1979:4 exhibits significantly more posterior uncertainty with the model at the mode being MA(1). The posterior probabilities, however, are virtually identical for the model trio AR(1), MA(1), and random walk.

The posteriors for both subsamples after the break points are very similar, with the mode at the AR(1) model and some posterior mass in the neighboring regions. The posterior uncertainty, however, is greater for the subsample starting in 1980.

The above lends credence to the interpretation that the random walk finding is at least to some extent driven by some large and persistent shift in the structure of the UK economy during the reign of Thatcher, consistent with conventional wisdom.

4.13.2 Impulse Responses

The impulse response functions are presented in Figure 4.18 and the estimation results from the frequentist regressions in Table 4.26. Again, the AICC and AIC tend to choose persistent models with oscillatory behavior, and the models chosen by the BIC are close to the responses from RJMCMC, except for the subsample starting in 1990. For both subsamples starting in 1960, the posterior of the impulse response is driven by random walk and low-order models. Both subsamples also show some extension of the credible sets into the negative after 1 quarter, in line with the impulse response of the models chosen by the frequentist criteria. The dominant model for both does not, however, exhibit meaningful persistence.

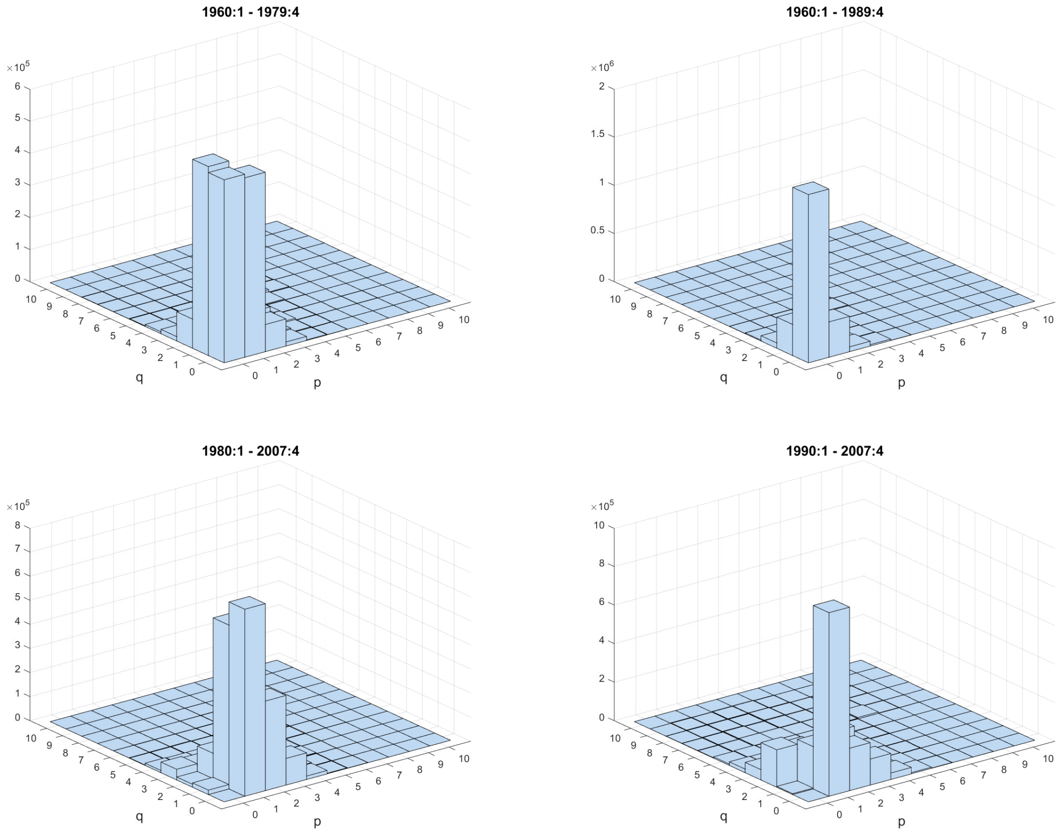

Figure 4.17: Posterior for model indicators for UK subsamples

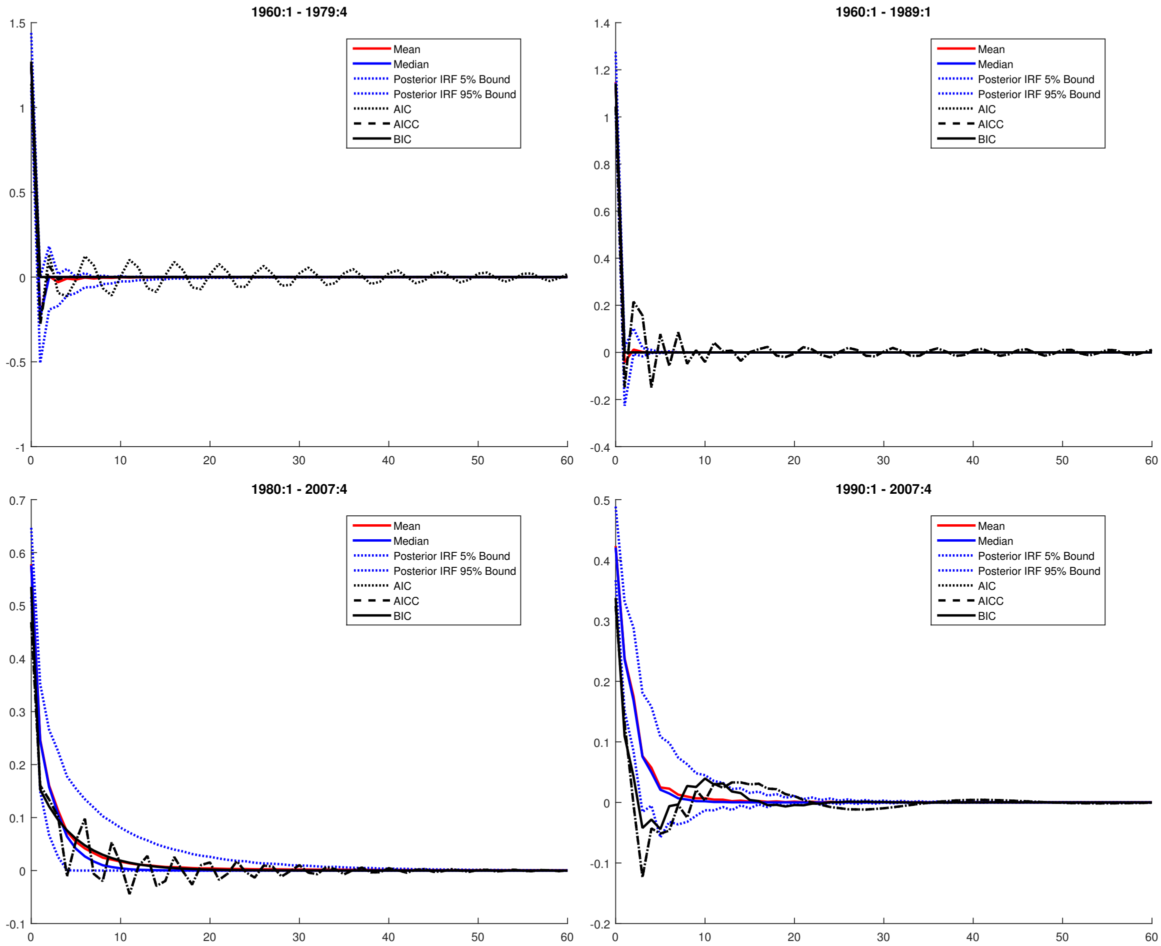

Figure 4.18: Estimated impulse responses for UK subsamples

The impulse response functions for both of the later subsamples show the familiar exponential decay of the response due to the preferred AR(1) model. For the subsample starting in 1980, the BIC chooses a model with a response virtually identical to the mean and mode responses from RJMCMC. While the credible sets are tight for both subsamples, the credible set for the subsample starting in 1990 includes some negative response after the third quarter. This negative response is also present in the reponses of the models chosen by all the frequentist criteria with the AICC and AIC again choosing a model with fairly persistent oscillatory behavior. Visually, the frequentist criteria seem to choose models with impulse responses at the borders of the credible sets from RJMCMC, roughly tracing out first the lower and then the upper bound.

Of note are also the magnitudes of the standard deviations. While the mean standard deviation for the first halves of the series is 1.259 for the series ending in 1979 and 1.147 for the subsample ending in 1989, the standard deviations for the second halves are significantly lower with 0.578 for the sample starting in 1980 and 0.423 for the one starting in 1990. This result is consistent with the standard deviation of the growth rates in the data: for the subsample ending in 1979 the standard deviation is 1.2778 and 1.1456 for the sample ending in 1989 while the standard deviations for the second subsamples are 0.6458 and 0.5176 respectively. This substantial shift in the variance of the growth rate is accompanied by the introduction of some persistence in the response of the growth rate to a shock, pointing towards something akin to a "great moderation", a phenomenon also seemingly present in US data. The question whether this diminished variance is due to successful economic policies reducing the variance of the shocks and/or smoothing their impact or simply luck has not been conclusively answered in the literature, neither for the UK nor the US, and it cannot be answered based on the results presented here.

Period	Criterion	P_1	P_2	P_3	P_4	Q_1	Q_2	Q_3	Q_4	Q_5	σ_e
1960:1 - 1979:4	AIC	-0.078	-0.575	-0.603		-0.164	0.673	0.391	-0.184		1.138
		(0.336)	(0.139)	(0.286)		(0.357)	(0.281)	(0.442)	(0.164)		(0.256)
	AICC	-0.229									1.238
		(0.114)									(0.165)
	BIC										1.270
											(0.181)
1980:1 - 2007:4	AIC	1.024	-0.766	1.000	-0.519	-0.681	0.701	-0.812	0.149	0.202	0.469
		(0.193)	(0.124)	(0.114)	(0.144)	(0.202)	(0.126)	(0.185)	(0.167)	(0.122)	(0.031)
	AICC	1.024	-0.766	1.000	-0.519	-0.681	0.701	-0.812	0.149	0.202	0.469
		(0.193)	(0.124)	(0.114)	(0.144)	(0.202)	(0.126)	(0.185)	(0.167)	(0.122)	(0.031)
	BIC	0.792				-0.506					0.535
		(0.048)				(0.104)					(0.032)
1960:1 - 1989:1	AIC	-0.242	-0.540	-0.759		0.100	0.712	0.883	-0.103	0.277	1.044
		(0.139)	(0.068)	(0.135)		(0.154)	(0.110)	(0.150)	(0.086)	(0.090)	(0.150)
	AICC	-0.242	-0.540	-0.759		0.100	0.712	0.883	-0.103	0.277	1.044
		(0.139)	(0.068)	(0.135)		(0.154)	(0.110)	(0.150)	(0.086)	(0.090)	(0.150)
	BIC										1.141
											(0.113)
1990:1 - 2007:4	AIC	0.343	1.163	-0.153	-0.525	0.063	-1.314	-0.697	0.600	0.542	0.324
		(0.087)	(0.094)	(0.111)	(0.089)	(0.150)	(0.136)	(0.212)	(0.166)	(0.110)	(0.022)
	AICC	0.343	1.163	-0.153	-0.525	0.063	-1.314	-0.697	0.600	0.542	0.324
		(0.087)	(0.094)	(0.111)	(0.089)	(0.150)	(0.136)	(0.212)	(0.166)	(0.110)	(0.022)
	BIC	1.364	0.145	-0.986	0.388	-1.040	-0.461	0.642			0.338
		(0.171)	(0.304)	(0.189)	(0.075)	(0.192)	(0.323)	(0.180)			(0.017)

Table 4.26: Frequentist regression results for UK subsamples

4.13.3 Persistence

The posterior distributions of the persistence measure presented in Figure 4.19 again reflect the behavior of the impulse responses. Point estimates from RJMCMC and frequentist estimation are also presented in the familiar form.

The large amount of posterior probability assigned to the random walk model is once more clearly visible through a clear mode at $C_n(1) = 1$ for the subsamples starting in 1960. Dispersion is somewhat greater for the shorter subsample. The posteriors for the subsamples starting in 1980 and 1990 resemble the posterior distributions for the other countries presented above in shape. The response of the growth rate to a disturbance is, however, quite persistent compared to the estimates for the other countries apart from Japan which cluster around a value of 1.5 whereas the point estimates for the later subsamples are 2.58 and 2.56 for the sample starting in 1980 and 1990 respectively. The UK would therefore not consistently be ranked in 6th place in terms of persistence but instead be second only to Japan.

Horizon	5	10	20	40	60
1960:1 - 1979:4	0.778 [0.804]	0.76 [0.804]	0.751 [0.804]	0.749 [0.804]	0.748 [0.804]
	[0.348; 1]	[0.208; 1]	[0.139; 1]	[0.117; 1]	[0.114; 1]
1980:1 - 2007:4	2.12 [2.08]	2.36 [2.21]	2.49 [2.23]	2.56 [2.23]	2.58 [2.23]
	[1.54; 2.86]	[1.53; 3.69]	[1.5; 4.4]	[1.5; 4.7]	[1.5; 4.74]
1960:1 - 1989:4	0.975 [1]	0.976 [1]	0.975 [1]	0.975 [1]	0.975 [1]
	[0.822; 1.06]	[0.818; 1.06]	[0.816; 1.06]	[0.815; 1.06]	[0.815; 1.06]
1990:1 - 2007:4	2.35 [2.28]	2.49 [2.34]	2.54 [2.34]	2.56 [2.34]	2.56 [2.34]
	[1.67; 3.27]	[1.59; 3.91]	[1.54; 4.22]	[1.52; 4.29]	[1.51; 4.3]

Table 4.27: $C_n(1)$ for UK subsamples at different horizons; RJMCMC estimates
Mean, [median] with [90% credible sets] in the second row

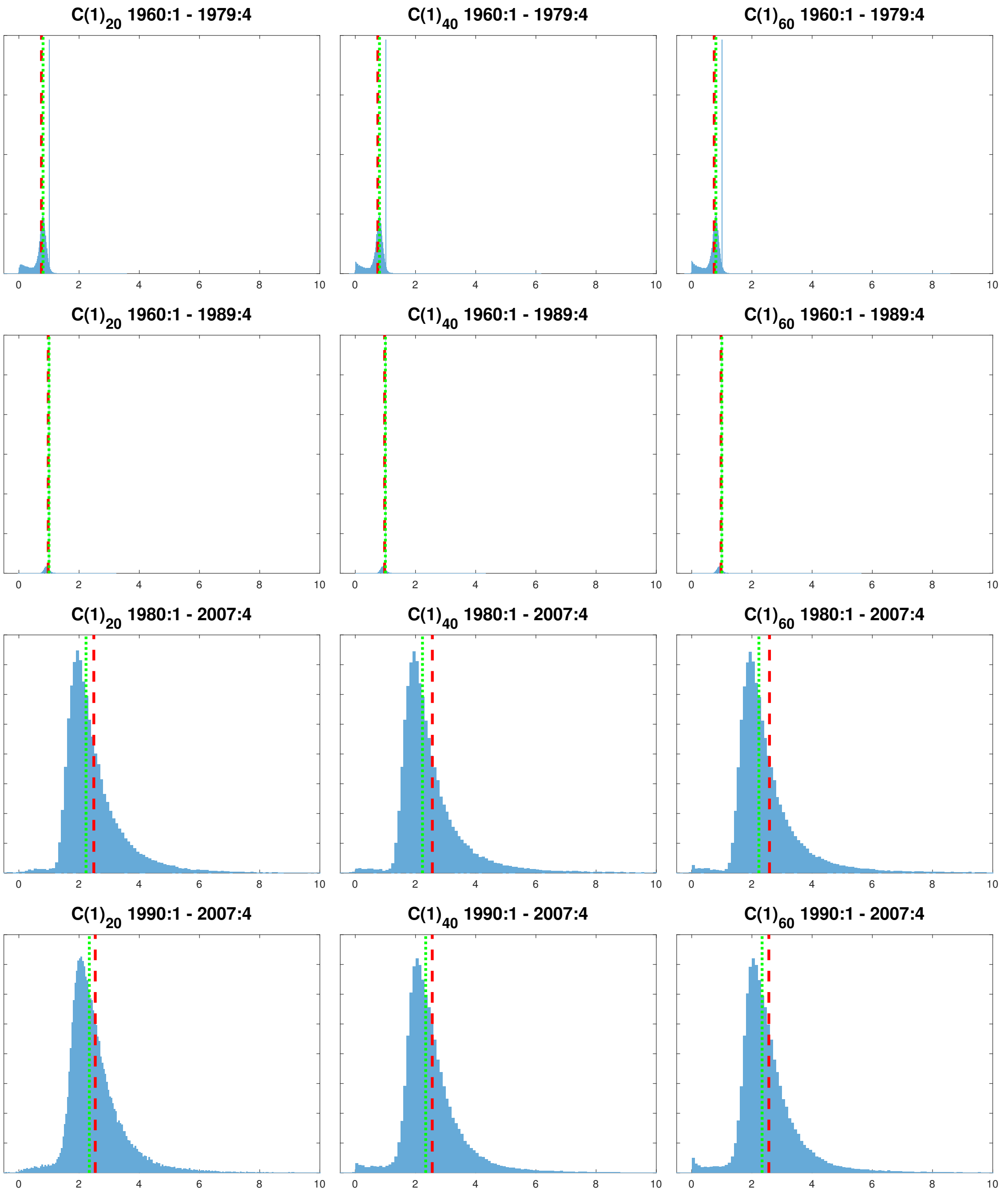

Figure 4.19: $C_n(1)$ for UK subsamples
Mean: Dashed line; Median: Dotted line

Horizon	5	10	20	40	60
1960:1 - 1979:4	$0.678;\ 1$	$0.697;\ 1$	$0.721;\ 1$	$0.749;\ 1$	$0.76;\ 1$
1980:1 - 2007:4	$1.94;\ 1.94$	$2.24;\ 2.24$	$2.16;\ 2.36$	$2.14;\ 2.37$	$2.14;\ 2.37$
1960:1 - 1989:4	$1.15;\ 1$	$1.1;\ 1$	$1.11;\ 1$	$1.12;\ 1$	$1.13;\ 1$
1990:1 - 2007:4	$0.722;\ 1.11$	$0.56;\ 1.36$	$1.33;\ 1.6$	$1.09;\ 1.59$	$1.12;\ 1.59$

Table 4.28: $C_n(1)$ for UK subsamples at different horizons
Frequentist estimates for AIC; BIC

4.13.4 Summary

The results presented above for the subsamples for UK GDP growth rates seem to support the conjecture that the random walk result for the whole series is driven by some large and persistent shifts in the level of GDP. When splitting the sample around the time of Margaret Thatcher, the random walk result only carries over for the first part of the series, while the following subsamples exhibit familiar patterns in terms of impulse responses as well as persistence with a drastically reduced variance of the disturbance. Whether this is a consequence of good policy or simple luck is unclear, but the dynamics of GDP do not seem to be constant over time, at least for the UK.

4.14 Conclusion

This chapter has investigated the dynamic behavior of real per capita GDP for six countries. Using a Bayesian approach, RJMCMC, posterior distributions accounting for model uncertainty have been obtained and analyzed using impulse response functions and a measure of persistence based on the infinite moving average representation of ARMA processes. The results have been compared to estimates obtained using maximum likelihood estimation while choosing a model according to three information criteria.

For all countries substantial persistence exists. Furthermore, strong differences in persistence across countries can be observed, with Japan being consistently ranked first in terms of persistence and exhibiting a degree of persistence far removed from the ones shown by the other economies analyzed. The results from frequentist estimates are mostly in line with the ones obtained using RJMCMC.

The estimates suggest that an innovation in the growth rate of GDP of 1% should induce an increased forecast for the level of GDP by substantially more than 1% in the future, consistent with results from other studies, most prominently the non-parametric estimates in Campbell and Mankiw (1989), with the sole exception of the UK. For this economy, the increase in the forecast should only be 1%, again roughly in line with the estimate from Campbell and Mankiw (1989) who also found the least persistence for the UK. This particular result is, however, sensitive to the time period studied. For example, using data starting in 1990, the corresponding increase in one's forecast for the level of GDP should be about 2.5%.

With regards to the ranking in terms of persistence across countries, the results presented here are mostly consistent with Campbell and Mankiw (1989). The behavior of the estimates as the horizon changes differs, however. While the estimates of Campbell and Mankiw (1989) increase with the horizon, RJMCMC estimates exhibit this pattern only for Japan and to some extent France. The magnitudes are also somewhat different, but the differences do not indicate a clear pattern.

The persistence ranking from a difference stationary perspective mostly carries over to OLS linear detrending, which has been used as a robustness check, offering only minor changes in the persistence ranking. The impulse responses are, however, significantly more persistent. These results contain a lesson for economic modeling: a model with a time trend must exhibit much stronger persistence in its impulse responses for output than a model featuring difference stationarity in order to capture the dynamics in the data.

Another robustness check was carried out using HP detrending. Here, the results appear to be dominated by filtering artifacts, casting doubt on the dependability of the estimates. Furthermore, it is questionable whether an analysis of long-run persistence is sensible when using a filter designed to extract a whole range of low frequencies from dynamics of the time series.

For the US, the dynamic behavior of the major components of GDP, private and government consumption, imports and exports, as well as fixed capital formation, were examined independently. The results for the aggregate series seem to be mainly driven by the two consumption series and to some extent by capital formation.

To conclude, while the question of difference vs trend stationarity could not be answered here, the results in this study suggest that significant persistence feature in the real GDP series for all countries studied. Shocks to GDP cast a long shadow into the future. The relative magnitude of persistence is robust to the detrending method, with the exception of the HP filter for which the estimates appear to be contaminated by filtering artifacts to a substantial degree. Persistence may, however, change over time as suggested by the results for subsamples for UK GDP.

4.15 Appendix

4.15.1 Additional Kolmogorov-Smirnov Results for First Differences

	Canada	France	Italy	Japan	UK	US
Canada	0	0.90783 (*)	0.2252 (*)	0.6005 (*)	0.926 (*)	0.29677 (*)
France	0.90783 (*)	0	0.94 (*)	0.98796 (*)	0.59668 (*)	0.93531 (*)
Italy	0.2252 (*)	0.94 (*)	0	0.44652 (*)	0.94552 (*)	0.09902 (*)
Japan	0.6005 (*)	0.98796 (*)	0.44652 (*)	0	0.98278 (*)	0.35091 (*)
UK	0.926 (*)	0.59668 (*)	0.94552 (*)	0.98278 (*)	0	0.93798 (*)
US	0.29677 (*)	0.93531 (*)	0.09902 (*)	0.35091 (*)	0.93798 (*)	0

Table 4.29: K-S test for $C(1)_5$ for first differences

	Canada	France	Italy	Japan	UK	US
Canada	0	0.64346 (*)	0.22286 (*)	0.84638 (*)	0.92358 (*)	0.24922 (*)
France	0.64346 (*)	0	0.73156 (*)	0.97455 (*)	0.43928 (*)	0.69557 (*)
Italy	0.22286 (*)	0.73156 (*)	0	0.79564 (*)	0.94258 (*)	0.06852 (*)
Japan	0.84638 (*)	0.97455 (*)	0.79564 (*)	0	0.99375 (*)	0.75236 (*)
UK	0.92358 (*)	0.43928 (*)	0.94258 (*)	0.99375 (*)	0	0.87506 (*)
US	0.24922 (*)	0.69557 (*)	0.06852 (*)	0.75236 (*)	0.87506 (*)	0

Table 4.30: K-S test for $C(1)_{10}$ for first differences

	Canada	France	Italy	Japan	UK	US
Canada	0	0.40579 (*)	0.22353 (*)	0.9368 (*)	0.92268 (*)	0.24155 (*)
France	0.40579 (*)	0	0.47572 (*)	0.94691 (*)	0.56207 (*)	0.42913 (*)
Italy	0.22353 (*)	0.47572 (*)	0	0.92163 (*)	0.94174 (*)	0.09432 (*)
Japan	0.9368 (*)	0.94691 (*)	0.92163 (*)	0	0.99661 (*)	0.91413 (*)
UK	0.92268 (*)	0.56207 (*)	0.94174 (*)	0.99661 (*)	0	0.85222 (*)
US	0.24155 (*)	0.42913 (*)	0.09432 (*)	0.91413 (*)	0.85222 (*)	0

Table 4.31: K-S test for $C(1)_{20}$ for first differences

	Canada	France	Italy	Japan	UK	US
Canada	0	0.36273 (*)	0.2237 (*)	0.94717 (*)	0.92238 (*)	0.24083 (*)
France	0.36273 (*)	0	0.41119 (*)	0.90935 (*)	0.58186 (*)	0.34913 (*)
Italy	0.2237 (*)	0.41119 (*)	0	0.93345 (*)	0.9414 (*)	0.09738 (*)
Japan	0.94717 (*)	0.90935 (*)	0.93345 (*)	0	0.99653 (*)	0.93455 (*)
UK	0.92238 (*)	0.58186 (*)	0.9414 (*)	0.99653 (*)	0	0.84894 (*)
US	0.24083 (*)	0.34913 (*)	0.09738 (*)	0.93455 (*)	0.84894 (*)	0

Table 4.32: K-S test for $C(1)_{30}$ for first differences

	Canada	France	Italy	Japan	UK	US
Canada	0	0.34851 (*)	0.22366 (*)	0.94887 (*)	0.92218 (*)	0.24043 (*)
France	0.34851 (*)	0	0.38823 (*)	0.84666 (*)	0.58882 (*)	0.3176 (*)
Italy	0.22366 (*)	0.38823 (*)	0	0.9329 (*)	0.94116 (*)	0.09906 (*)
Japan	0.94887 (*)	0.84666 (*)	0.9329 (*)	0	0.99629 (*)	0.93882 (*)
UK	0.92218 (*)	0.58882 (*)	0.94116 (*)	0.99629 (*)	0	0.84777 (*)
US	0.24043 (*)	0.3176 (*)	0.09906 (*)	0.93882 (*)	0.84777 (*)	0

Table 4.33: K-S test for $C(1)_{50}$ for first differences

	Canada	France	Italy	Japan	UK	US
Canada	0	0.34698 (*)	0.22366 (*)	0.94891 (*)	0.92212 (*)	0.24039 (*)
France	0.34698 (*)	0	0.38588 (*)	0.82598 (*)	0.58954 (*)	0.31435 (*)
Italy	0.22366 (*)	0.38588 (*)	0	0.93246 (*)	0.94109 (*)	0.09942 (*)
Japan	0.94891 (*)	0.82598 (*)	0.93246 (*)	0	0.99605 (*)	0.93945 (*)
UK	0.92212 (*)	0.58954 (*)	0.94109 (*)	0.99605 (*)	0	0.84764 (*)
US	0.24039 (*)	0.31435 (*)	0.09942 (*)	0.93945 (*)	0.84764 (*)	0

Table 4.34: K-S test for $C(1)_{60}$ for first differences

4.15.2 Additional Kolmogorov-Smirnov Results for OLS detrended Data

	Canada	France	Italy	Japan	UK	US
Canada	0	0.98085 (*)	0.26592 (*)	0.46009 (*)	0.95862 (*)	0.02384 (*)
France	0.98085 (*)	0	0.99191 (*)	0.99601 (*)	0.34982 (*)	0.97951 (*)
Italy	0.26592 (*)	0.99191 (*)	0	0.28987 (*)	0.97918 (*)	0.26564 (*)
Japan	0.46009 (*)	0.99601 (*)	0.28987 (*)	0	0.98778 (*)	0.48219 (*)
UK	0.95862 (*)	0.34982 (*)	0.97918 (*)	0.98778 (*)	0	0.95492 (*)
US	0.02384 (*)	0.97951 (*)	0.26564 (*)	0.48219 (*)	0.95492 (*)	0

Table 4.35: K-S test for $C(1)_5$ for linear trend

	Canada	France	Italy	Japan	UK	US
Canada	0	0.85326 (*)	0.33398 (*)	0.77479 (*)	0.96433 (*)	0.36318 (*)
France	0.85326 (*)	0	0.94166 (*)	0.98908 (*)	0.37822 (*)	0.62025 (*)
Italy	0.33398 (*)	0.94166 (*)	0	0.6546 (*)	0.98744 (*)	0.63403 (*)
Japan	0.77479 (*)	0.98908 (*)	0.6546 (*)	0	0.99753 (*)	0.8968 (*)
UK	0.96433 (*)	0.37822 (*)	0.98744 (*)	0.99753 (*)	0	0.85126 (*)
US	0.36318 (*)	0.62025 (*)	0.63403 (*)	0.8968 (*)	0.85126 (*)	0

Table 4.36: K-S test for $C(1)_{10}$ for linear trend

	Canada	France	Italy	Japan	UK	US
Canada	0	0.50011 (*)	0.46724 (*)	0.93963 (*)	0.95242 (*)	0.76679 (*)
France	0.50011 (*)	0	0.7902 (*)	0.98207 (*)	0.73116 (*)	0.38668 (*)
Italy	0.46724 (*)	0.7902 (*)	0	0.87262 (*)	0.99439 (*)	0.91873 (*)
Japan	0.93963 (*)	0.98207 (*)	0.87262 (*)	0	0.99981 (*)	0.99221 (*)
UK	0.95242 (*)	0.73116 (*)	0.99439 (*)	0.99981 (*)	0	0.3832 (*)
US	0.76679 (*)	0.38668 (*)	0.91873 (*)	0.99221 (*)	0.3832 (*)	0

Table 4.37: K-S test for $C(1)_{20}$ for linear trend

	Canada	France	Italy	Japan	UK	US
Canada	0	0.27342 (*)	0.54445 (*)	0.96783 (*)	0.92406 (*)	0.84404 (*)
France	0.27342 (*)	0	0.68925 (*)	0.97414 (*)	0.8345 (*)	0.7238 (*)
Italy	0.54445 (*)	0.68925 (*)	0	0.91659 (*)	0.99201 (*)	0.96143 (*)
Japan	0.96783 (*)	0.97414 (*)	0.91659 (*)	0	0.99983 (*)	0.9971 (*)
UK	0.92406 (*)	0.8345 (*)	0.99201 (*)	0.99983 (*)	0	0.13508 (*)
US	0.84404 (*)	0.7238 (*)	0.96143 (*)	0.9971 (*)	0.13508 (*)	0

Table 4.38: K-S test for $C(1)_{30}$ for linear trend

	Canada	France	Italy	Japan	UK	US
Canada	0	0.06116 (*)	0.59258 (*)	0.95936 (*)	0.87519 (*)	0.86943 (*)
France	0.06116 (*)	0	0.61963 (*)	0.94099 (*)	0.87797 (*)	0.8719 (*)
Italy	0.59258 (*)	0.61963 (*)	0	0.89009 (*)	0.98216 (*)	0.97407 (*)
Japan	0.95936 (*)	0.94099 (*)	0.89009 (*)	0	0.99938 (*)	0.99674 (*)
UK	0.87519 (*)	0.87797 (*)	0.98216 (*)	0.99938 (*)	0	0.02389 (*)
US	0.86943 (*)	0.8719 (*)	0.97407 (*)	0.99674 (*)	0.02389 (*)	0

Table 4.39: K-S test for $C(1)_{50}$ for linear trend

	Canada	France	Italy	Japan	UK	US
Canada	0	0.03504 (*)	0.59639 (*)	0.93725 (*)	0.85954 (*)	0.87205 (*)
France	0.03504 (*)	0	0.61165 (*)	0.91461 (*)	0.86784 (*)	0.87745 (*)
Italy	0.59639 (*)	0.61165 (*)	0	0.84468 (*)	0.97737 (*)	0.97514 (*)
Japan	0.93725 (*)	0.91461 (*)	0.84468 (*)	0	0.99751 (*)	0.99421 (*)
UK	0.85954 (*)	0.86784 (*)	0.97737 (*)	0.99751 (*)	0	0.04726 (*)
US	0.87205 (*)	0.87745 (*)	0.97514 (*)	0.99421 (*)	0.04726 (*)	0

Table 4.40: K-S test for $C(1)_{60}$ for linear trend

4.15.3 Additional Kolmogorov-Smirnov Results for HP detrended Data

	Canada	France	Italy	Japan	UK	US
Canada	0	0.96313 (*)	0.55185 (*)	0.31978 (*)	0.55247 (*)	0.17481 (*)
France	0.96313 (*)	0	0.83271 (*)	0.91359 (*)	0.87615 (*)	0.98294 (*)
Italy	0.55185 (*)	0.83271 (*)	0	0.25863 (*)	0.06192 (*)	0.68343 (*)
Japan	0.31978 (*)	0.91359 (*)	0.25863 (*)	0	0.24069 (*)	0.47301 (*)
UK	0.55247 (*)	0.87615 (*)	0.06192 (*)	0.24069 (*)	0	0.68844 (*)
US	0.17481 (*)	0.98294 (*)	0.68343 (*)	0.47301 (*)	0.68844 (*)	0

Table 4.41: K-S test for $C(1)_5$ for HP filter

	Canada	France	Italy	Japan	UK	US
Canada	0	0.39382 (*)	0.50325 (*)	0.08816 (*)	0.19114 (*)	0.12875 (*)
France	0.39382 (*)	0	0.22189 (*)	0.47737 (*)	0.25589 (*)	0.27229 (*)
Italy	0.50325 (*)	0.22189 (*)	0	0.58397 (*)	0.39561 (*)	0.39193 (*)
Japan	0.08816 (*)	0.47737 (*)	0.58397 (*)	0	0.27676 (*)	0.2166 (*)
UK	0.19114 (*)	0.25589 (*)	0.39561 (*)	0.27676 (*)	0	0.07209 (*)
US	0.12875 (*)	0.27229 (*)	0.39193 (*)	0.2166 (*)	0.07209 (*)	0

Table 4.42: K-S test for $C(1)_{10}$ for HP filter

	Canada	France	Italy	Japan	UK	US
Canada	0	0.33491 (*)	0.40334 (*)	0.4021 (*)	0.07921 (*)	0.17017 (*)
France	0.33491 (*)	0	0.12849 (*)	0.12412 (*)	0.32156 (*)	0.4936 (*)
Italy	0.40334 (*)	0.12849 (*)	0	0.18634 (*)	0.39056 (*)	0.54647 (*)
Japan	0.4021 (*)	0.12412 (*)	0.18634 (*)	0	0.41018 (*)	0.53602 (*)
UK	0.07921 (*)	0.32156 (*)	0.39056 (*)	0.41018 (*)	0	0.24275 (*)
US	0.17017 (*)	0.4936 (*)	0.54647 (*)	0.53602 (*)	0.24275 (*)	0

Table 4.43: K-S test for $C(1)_{20}$ for HP filter

	Canada	France	Italy	Japan	UK	US
Canada	0	0.29307 (*)	0.16128 (*)	0.19692 (*)	0.05716 (*)	0.27729 (*)
France	0.29307 (*)	0	0.2301 (*)	0.09745 (*)	0.28851 (*)	0.52728 (*)
Italy	0.16128 (*)	0.2301 (*)	0	0.15904 (*)	0.12747 (*)	0.43728 (*)
Japan	0.19692 (*)	0.09745 (*)	0.15904 (*)	0	0.19531 (*)	0.43287 (*)
UK	0.05716 (*)	0.28851 (*)	0.12747 (*)	0.19531 (*)	0	0.32595 (*)
US	0.27729 (*)	0.52728 (*)	0.43728 (*)	0.43287 (*)	0.32595 (*)	0

Table 4.44: K-S test for $C(1)_{30}$ for HP filter

	Canada	France	Italy	Japan	UK	US
Canada	0	0.25295 (*)	0.08726 (*)	0.19597 (*)	0.02858 (*)	0.25459 (*)
France	0.25295 (*)	0	0.33926 (*)	0.1054 (*)	0.23759 (*)	0.47643 (*)
Italy	0.08726 (*)	0.33926 (*)	0	0.24071 (*)	0.10391 (*)	0.28712 (*)
Japan	0.19597 (*)	0.1054 (*)	0.24071 (*)	0	0.17015 (*)	0.44848 (*)
UK	0.02858 (*)	0.23759 (*)	0.10391 (*)	0.17015 (*)	0	0.28146 (*)
US	0.25459 (*)	0.47643 (*)	0.28712 (*)	0.44848 (*)	0.28146 (*)	0

Table 4.45: K-S test for $C(1)_{40}$ for HP filter

	Canada	France	Italy	Japan	UK	US
Canada	0	0.13596 (*)	0.17464 (*)	0.13558 (*)	0.03753 (*)	0.13773 (*)
France	0.13596 (*)	0	0.14547 (*)	0.10886 (*)	0.14447 (*)	0.1346 (*)
Italy	0.17464 (*)	0.14547 (*)	0	0.12736 (*)	0.17647 (*)	0.12889 (*)
Japan	0.13558 (*)	0.10886 (*)	0.12736 (*)	0	0.15701 (*)	0.10085 (*)
UK	0.03753 (*)	0.14447 (*)	0.17647 (*)	0.15701 (*)	0	0.15547 (*)
US	0.13773 (*)	0.1346 (*)	0.12889 (*)	0.10085 (*)	0.15547 (*)	0

Table 4.46: K-S test for $C(1)_{50}$ for HP filter

	Canada	France	Italy	Japan	UK	US
Canada	0	0.13528 (*)	0.06751 (*)	0.10346 (*)	0.03713 (*)	0.06562 (*)
France	0.13528 (*)	0	0.12716 (*)	0.1284 (*)	0.14378 (*)	0.13572 (*)
Italy	0.06751 (*)	0.12716 (*)	0	0.14865 (*)	0.10131 (*)	0.10688 (*)
Japan	0.10346 (*)	0.1284 (*)	0.14865 (*)	0	0.12635 (*)	0.103 (*)
UK	0.03713 (*)	0.14378 (*)	0.10131 (*)	0.12635 (*)	0	0.08265 (*)
US	0.06562 (*)	0.13572 (*)	0.10688 (*)	0.103 (*)	0.08265 (*)	0

Table 4.47: K-S test for $C(1)_{60}$ for HP filter

Chapter 5

Reversible Jump Markov Chain Monte Carlo vs. Frequentist Information Criteria: A Horse Race

Reversible Jump Markov Chain Monte Carlo vs. Frequentist Information Criteria
A Horse Race *

Daniel Neuhoff †

Abstract

This study investigates the performance of Reversible Jump Markov Chain Monte Carlo for the estimation of autoregressive moving average models of unknown order compared to maximum likelihood estimates of the same, with model choice according to the Akaike Information Criterion, the Corrected Akaike Information Criterion, as well as the Bayesian Information Criterion. The performance of the approaches is compared in terms of model selection, as well as their ability to match the true impulse response functions. Reversible Jump Markov Chain Monte Carlo outperforms the alternative approaches in almost all experiments, closely followed by the Bayesian Information Criterion which exhibits very similar performance in some cases. Both the Akaike and the Corrected Akaike Information Criteria perform substantially worse. Furthermore, the influence of a break in a deterministic time trend on estimates of the persistence measure from chapter 4 is studied. The results indicate, that tests of trend vs difference stationary hypotheses based on aforementioned persistence measure cannot identify trend stationarity in this setup.

JEL classification: C11, C52

Keywords: ARMA; Persistence; Reversible Jump Markov Chain Monte Carlo; Stationarity; Non-Normality

---

*I am grateful to Michael Burda and Alexander Meyer-Gohde for useful comments. Furthermore, I am indebted to Tobias König for excellent research assistance. This research was supported by the DFG through the SFB 649 "Economic Risk." Any and all errors are entirely my own.

†Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5667; E-Mail:daniel.neuhoff@gmail.com

5.1 Introduction

This study investigates the performance of the Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm for the estimation of autoregressive moving average (ARMA) models employed in chapters 2 and 4 of this thesis in a variety of settings and compares the results to those obtained using classical estimation and model selection techniques. The classical or frequentist approach employed here is maximum likelihood estimation combined with model selection using three information criteria, the Akaike Information Criterion, the Corrected Akaike Information Criterion, and the Schwartz or Bayesian Information Criterion. The estimates are compared on the criteria of model selection and identification of impulse response functions.

Several experiments were carried out to shed light on these questions. The first experiment compares the performance of the approaches if the likelihood function is misspecified. For this experiment, the disturbances for the models were generated using a truncated Cauchy distribution as well as a mixture of two normal distributions. The next experiment analyzes the influence of sample sizes. The last experiment aims to characterize the effects of failing to account for a break in a linear time trend when detrending using OLS, and studies the ability of estimates of the persistence measure from Campbell and Mankiw (1987), also employed in chapter 4, to identify trend stationarity when the data is first differenced in this context.

The remainder of this chapter is structured as follows. After a short primer on Reversible Jump Markov Chain Monte Carlo and a short description of the implementation of the frequentist regressions, the two criteria used for gauging the performance of the methods are explained. This sets up a short section on data generation. Then, the two experiments regarding the misspecified likelihood function are presented. Afterwards, the influence of the sample size on the performance of the different approaches is studied, leading to the analysis regarding a break in a linear trend. The last section concludes.

5.2 Reversible Jump Markov Chain Monte Carlo

The estimation approach to be tested in this study, Reversible Jump Markov Chain Monte Carlo (RJMCMC), was pioneered by Green (1995). It enables sampling from a posterior distribution spanning several, not necessarily nested, models where the dimensionality of the parameter space associated with these models may vary.

The method as applied here generates posterior distributions spanning the model and corresponding parameter spaces of stationary ARMA(p,q) models of the form:

$$P(L)y_t = Q(L)\epsilon_t; \epsilon_t \sim N(0, \sigma_e^2) \tag{5.1}$$

with

$$P(L) = 1 - P_1 L - P_2 L^2 - ...P_p L^p \tag{5.2}$$

$$Q(L) = 1 + Q_1 L + Q_2 L^2 + ...Q_q L^q \tag{5.3}$$

denoting the autoregressive and moving average polynomials respectively and L denoting the lag operator. In particular, it is assumed throughout that the coefficients of $Q(L)$ satisfy the invertibility and those of $P(L)$ the stationarity conditions, and $p, q \in [0; 10]$.[1] In order to impose these conditions, the model is reparametrized in terms of the (inverse) partial autocorrelations for the (moving average) autoregressive polynomials as in e.g.Barndorff-Nielsen and Schou (1973), Monahan (1984), Jones (1987), and chapters 2 as well as 4 of this thesis.

The implementation employed in this study is identical to the one used in chapter 4, including the adoption of different proposal distributions for between- and within-model moves and the evaluation of the likelihood by means of the Kalman filter. A more in-depth explanation of the algorithm as well as further literature can be found in chapter 2. Among the literature

[1] Invertibility of the moving average polynomial is commonly assumed in order to ensure identification. The assumption of stationarity represents a modeling choice, corresponding to the one taken in the other chapters of this thesis. There, once the data is detrended using first differences, a linear trend, or the Hodrick-Prescott filter, the remainder is assumed to be stationary.

cited there, Waagepetersen and Sorensen (2001) deserves special mention as it provides an excellent tutorial for the construction of an RJMCMC sampler.

The prior structure applied here is the same as in chapters 2 and 4 and assumes a priori independence of all parameters. The priors reported in Table 5.1 are the same for all variants considered. In Table 5.1, $DU(a, b)$ denotes the discrete uniform distribution on the interval

Object	Prior
p	$DU(0, 10)$
q	$DU(0, 10)$
(Inverse) Partial Autocorrelation	$U(-1, 1)$
σ_ϵ	$IG(1, 1)$

Table 5.1: Prior distributions

$[a, b]$, $U(a, b)$ is the continuous uniform distribution on the open interval (a, b) and $IG(a, b)$ denotes the inverse gamma distribution with parameters (a, b). p and q are the orders of the autoregressive and moving average lag polynomials respectively, and σ_ϵ is the standard deviation of the disturbances ϵ_t. Further discussion of the prior structure can be found in chapters 2 and 4.

The proposal distributions and their parameters are presented in the following sections. The proposals were tuned in short pilot runs by tweaking the standard deviation of the proposal distributions for each experiment, targeting acceptance rates of 10% for between model moves and 30% for within model moves.[2]

[2] Between model moves are those for which the order of at least one lag polynomial changes while for within model moves the polynomial orders remain constant. The acceptance rate gives the share of proposed moves that are accepted during sampling. Acceptance rates will be reported both averaged over the two move types, as well as individually.

5.3 Frequentist Regressions

The frequentist estimates were obtained using the same approaches and tools as in chapter 4. That is, the maximum likelihood estimations were carried out using the Econometrics Toolbox of Matlab 2015a. For the frequentist estimates, the model space was constrained to include only models with autoregressive and moving average lag polynomials up to degree five, i.e. $p \leq 5 \wedge q \leq 5$.[3]

In order to pick a model, three information criteria were employed: The Akaike Information Criterion (AIC), see Akaike (1974), the Corrected Akaike Information Criterion (AICC), see Sugiura (1978) and Hurvich and Tsai (1989), and the Schwarz or Bayesian Information Criterion (BIC), see Schwarz (1978). These are given by:

$$\mathrm{AIC} = 2k - 2\ln(\hat{\mathcal{L}}), \quad \mathrm{AICC} = AIC + \frac{2k(k+1)}{n-k-1}, \quad \mathrm{BIC} = k\ln(n) - 2\ln(\hat{\mathcal{L}})$$

with k being the number of model parameters and n the number of observations. $\hat{\mathcal{L}}$ denotes the maximized likelihood value of a model, i.e., for given ARMA orders p and q. The model chosen is then the one with the lowest value of the information criterion which is being applied.[4]

The AIC is known to have a strong tendency of overestimating the number of parameters when the true model has a finite number of parameters as noted in Hurvich and Tsai (1989). They encourage the use of the corrected Akaike Information Criterion (AICC), especially for situations in which the sample size is small. In the experiments presented here, however, the AICC does not perform better than the AIC.

[3] Many authors restrict the model space even further, e.g. to $p + q \leq 6$ as e.g. in Diebold and Rudebusch (1989). The truncation of the model space chosen here is the same as in Perron (1993).

[4] Both the AIC and AICC are derived as estimates of the expected Kullback-Leibler information and model selection is based on minimizing Kullback's discrimination information, see Sugiura (1978). In contrast, model selection using the BIC amounts to picking the model with highest posterior probability, see Schwarz (1978). See e.g. Chow (1981) for a comparison.

While the AIC and AICC are asymptotically efficient in terms of minimizing the mean squared error of prediction in a situation where the true model is of infinite order, the BIC is consistent in the sense that it asymptotically chooses the correct model with probability one if the true model is of finite dimension and included in the set of models under consideration. The results regarding the ability of the methods to identify the true model presented below are then not surprising with the BIC consistently choosing more parsimonious models compared to the other two criteria.

The tendency towards overfitting of the Akaike criteria need not necessarily impact their ability to match the true impulse responses since higher-order ARMA models may exhibit very similar dynamic properties compared to a lower-order one due to root cancellation discussed below. The results show, however, that the BIC and RJMCMC also outperform the AIC and AICC in this respect in the setups considered here.

5.4 Modes of Comparison

In this study, the performance of the approaches is compared along different dimensions. First, the share of correctly identified models for both the frequentist information criteria as well as RJMCMC are presented. The model choice for RJMCMC was the model at the mode in (p, q), thus the one with the highest posterior probability. This setup is less than ideal for RJMCMC, since posterior model uncertainty may be substantial, particularly due to the reasons discussed in the following, in which case RJMCMC would not necessarily produce a posterior distribution with the mode at the right model when posterior model probabilities are very similar.

ARMA models with high but different orders can possess very similar autocorrelation structures, which in turn define the likelihood, resulting in difficulties regarding model identification. Take, for example, the extreme case in which the autoregressive and moving average lag polynomials of order p and q respectively, share a common root. In this case, this ARMA(p,q) model is the same as the ARMA(p-1,q-1) model. This can be easily seen by way of the following argument.

Let again $P(L)$ be the autoregressive and $Q(L)$ be the moving average lag polynomials. Factoring the lag polynomials in terms of their roots λ_i^P and λ_j^Q gives:

$$P(L)y_t = Q(L)\epsilon_t$$
$$(1-\lambda_1^P L)(1-\lambda_2^P L)\ldots(1-\lambda_p^P L)y_t = (1-\lambda_1^Q L)(1-\lambda_2^Q L)\ldots(1-\lambda_q^Q L)\epsilon_t$$

If any two roots λ_i^P and λ_j^Q have the same value, the orders of the polynomials on both sides can obviously be reduced by one without changing the model. This phenomenon is well known under the name root-cancellation, and poses a problem for the selection of ARMA(p,q) models.

Thus, the performance of the procedures was also compared in terms of their ability to match the true impulse response functions which capture the dynamics of the model. The distance measure employed here is the sum of the squared distances of the estimated and true impulse responses to a one standard deviation shock up to different horizons. The ability of the methods to match the standard deviation of the disturbance is thus also carries weight in the measure. For RJMCMC output, this comparison is based on the mean and median responses at each horizon. The mean and median estimates were based on every 30th draw from the posterior in order to keep computation times manageable.

Let Φ_k be the true impulse response at horizon k and $\Phi_k(\hat{\beta})$ the impulse response implied by the estimates $\hat{\beta}$.[5] The distance measure for horizon i, D_i, is then given by

$$D_i = \sum_{k=0}^{i} \left(\Phi_k - \Phi_k(\hat{\beta})\right)^2$$

amounting to a quadratic loss function for impulse responses. By way of this comparison, the ability of the approaches to pin down the dynamics of the time series is characterized

[5] The term estimate is meant to include the model chosen.

well using this one-dimensional statistic. Furthermore, this statistic is a univariate version of the objective function of estimators which are based on matching empirical impulse response function from vector autoregressions as employed for example in Christiano, Eichenbaum, and Evans (2005).

5.5 Data Generation

The synthetic data for this study were generated from zero-mean and stationary ARMA processes with initial values set to zero, i.e. the steady state of the process. Then, $n = (1000 + samplesize)$ realizations of the disturbances were sampled from the given distribution of the error terms and the corresponding number of observations of the process calculated by iterating forward. Finally, the first 1000 observations were discarded in order to minimize the effect of the initial state.

For each experiment in which the model is not fixed across data sets, the lag polynomial orders were drawn independently from the discrete uniform distribution over the interval $[0, 5]$. The corresponding model parameters were generated by first drawing the appropriate number of (inverse) partial autocorrelations from integer-parameter beta distributions $Beta([\frac{1}{2}(k+1)], [\frac{1}{2}k] + 1)$ with k the order of the partial autocorrelation, so as to obtain parameters which are uniformly distributed over the invertibility and stationarity region as described in Jones (1987).[6]

5.6 Misspecified Likelihood Function

This section will investigate the influence of misspecification of the likelihood function. With real world data, the assumption of normally distributed disturbances does not necessarily have

[6] That is, the first partial autocorrelation for e.g. an autoregressive polynomial is drawn from the Beta(1,1) distribution, the second from Beta(1,2), the third from Beta(2,2), the fourth from Beta(2,3), and so forth until the order of the polynomial is reached.

to be correct. That is, while the likelihood function is assumed to be standard Gaussian, the actual distribution of the disturbances in this section is a mixture of two normal distributions in the first experiment, as well as a truncated Cauchy distribution in the second.

For each of the two experiments, 100 data sets with 250 observations each were generated, with the model and parameters sampled as described in the foregoing. For each data set in this and all following experiments, 1,500,000 samples from the posterior were obtained, discarding the first 1,000,000 samples as burn-in.

Optimally, the sampler settings would be tuned by tweaking the standard deviation of the proposal distributions for each data set, but this approach is wrought with a prohibitive penalty regarding computational effort for the number of data sets considered in this study. For some data sets, no between-model moves were accepted and the chains would have been re-run with different sampler settings when studying a smaller number of data sets. The results presented here can thus be interpreted as a lower bound on performance which would be improved by increasing the amount of samples from the posterior, tuning the sampler settings for each data set and/or using an adaptive approach for the proposals as e.g. in Ehlers and Brooks (2008). The theory and practice of adaptive samplers, however, are still in their infancy and pilot tuning remains the most widely used approach when applying RJMCMC in practice.

5.6.1 Mixture of Normals

This section analyzes the performance of the methods when the disturbances are sampled from a mixture of two normal distributions with means 0.75 and -0.75 and standard deviations equal to 0.4. The lag polynomial orders and parameters were generated according to the procedure laid out in the foregoing.

Table 5.2 reports the proposal distributions and their parameters employed in this experiment. Here, $DL(\mu, b)$ denotes the discretized Laplace distribution, with location parameter, μ, and shape parameter, b, such that $P(x|\mu) \propto exp(-b|\mu - x|) \text{ with } \mu, x \in [0, 1, \ldots, 10]$. $TN(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2 truncated to the interval $(-1, 1)$ for the partial autocorrelations and $(0, 1000)$ for the standard deviation of the error term σ_ϵ respectively.

Object	Proposal
p	$DL(\mu, 2.2)$
q	$DL(\mu, 2.2)$
(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.045^2)$
(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.02^2)$
σ_ϵ	$TN(\mu, 0.04^2)$

Table 5.2: Proposal distributions for mixture of normals
$DL(\mu, b)$ denotes the discretized Laplace distribution; $TN(\mu, \sigma^2)$ denotes the truncated normal distribution

Table 5.3 presents summary statistics for the acceptance rates and number of model parameters across the 100 data sets and corresponding models. The average overall acceptance rate (α) achieved was 29.31 %, with averages for within- (α_w) and between- (α_b) model moves of 42% and 5% respectively. The low acceptance rates for some data sets are due to some models with a large number of parameters. The correlation of the sum of the true lag polynomial orders $p+q$ with the acceptance rate for within (between) model moves was -0.585 (-0.657).[7]

	True p	True q	α	α_w	α_b
Mean	2.54	2.74	0.29	0.42	0.05
Standard Deviation	1.6662	1.6552	0.11	0.14	0.02
Range	[0; 5]	[0; 5]	[0.04; 0.52]	[0.07; 0.66]	[0.00; 0.13]

Table 5.3: Acceptance rates for mixture of normals
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

Table 5.4 reports the shares of correctly identified models for the four approaches. RJM-

[7] Some of the chains would have been discarded and re-run with different settings. By tuning the sampler for each data set, significantly better performance can be expected and the results presented here represent a lower bound on performance, as argued above.

CMC clearly outperforms the other criteria in this setup, correctly identifying the model in 35% of the samples. The Akaike criteria perform substantially worse, identifying the correct model only for 15 out of 100 data sets. The BIC is in second place with a share of correctly identified models of 0.28.

Criterion	Correct Identification
RJMCMC	0.33
AIC	0.15
AICC	0.15
BIC	0.28

Table 5.4: Share of correctly identified models for mixture of normals

Table 5.5 presents several statistics on D_i,[8] namely the average size of D_i across data sets and models, the median in square brackets, as well as the standard deviation in round brackets. Clearly, RJMCMC not only outperforms the frequentist approaches in this experiment with respect to the model choice, but also with respect to matching the true impulse response function with smaller mean and median mismatch lower variability of the measure across data sets.

To conclude, at least in this setup, RJMCMC offers substantially better performance with respect to matching the true impulse response at all horizons considered, as well as model selection.

[8] Let Φ_k be the true impulse response at horizon k and $\Phi_k(\hat{\beta})$ the impulse response implied by the estimates $\hat{\beta}$. The distance measure for horizon i, D_i, is then given by $D_i = \sum_{k=0}^{i} \left(\Phi_k - \Phi_k(\hat{\beta}))^2 \right)$ with $i \in \{5, 10, 20, 40, 60\}$.

Horizon	5	10	20	40	60
RJMCMC Mean	0.0539 [0.0173] (0.166)	0.273 [0.0354] (1.69)	0.82 [0.0498] (5.36)	1.34 [0.059] (8.23)	1.58 [0.0615] (9.74)
RJMCMC Median	0.0565 [0.0166] (0.187)	0.288 [0.0346] (1.8)	0.812 [0.0476] (5.42)	1.19 [0.0516] (7.12)	1.32 [0.0516] (7.65)
AIC	0.111 [0.0257] (0.511)	0.464 [0.057] (2.37)	1.28 [0.0945] (8.14)	1.86 [0.137] (10.8)	2.07 [0.15] (11.5)
AICC	0.119 [0.0265] (0.587)	0.479 [0.0555] (2.47)	1.28 [0.0945] (8.14)	1.81 [0.137] (10.5)	2.01 [0.15] (11.1)
BIC	0.123 [0.0228] (0.591)	0.566 [0.0489] (3.36)	1.28 [0.0699] (8.58)	1.78 [0.102] (10.7)	1.95 [0.105] (11.4)

Table 5.5: D_i for mixture of normals at different horizons
Mean, [Median], and (Standard Deviation)

5.6.2 Cauchy Distribution

This section analyzes the performance of the methods when the disturbances are sampled from a truncated Cauchy distribution. The location parameter was set to zero and the shape parameter to 0.5. The distribution was truncated to lie on the interval $[-3, 3]$. The lag polynomial orders and parameters were again generated following the procedure laid out in the foregoing.

Table 5.6 reports the proposal distributions and their parameters employed for this experiment.[9]

Object	Proposal
p	$DL(\mu, 2.2)$
q	$DL(\mu, 2.2)$
(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.03^2)$
σ_ϵ	$TN(\mu, 0.04^2)$

Table 5.6: Proposal distributions for Cauchy distribution
$DL(\mu, b)$ denotes the discretized Laplace distribution; $TN(\mu, \sigma^2)$ denotes the truncated normal distribution

Table 5.7 presents summary statistics for the models and acceptance rates. The average overall acceptance rate achieved was 24.7%, with averages for within and between model moves of 34% and 5% respectively. The low acceptance rates for between-model moves are again due to models with a large number of parameters. The correlation of $p + q$ with the acceptance rate for within (between) model moves was -0.662 (-0.694). The arguments in the foregoing regarding the acceptance rates and validity of the results apply here as well.

The results for Cauchy distributed disturbances show a somewhat different picture than for the experiment before. While RJMCMC exhibits performance similar to the foregoing experi-

[9] $DL(\mu, b)$ denotes the discretized Laplace distribution, with location parameter, μ, and shape parameter, b, such that $P(x|\mu) \propto exp(-b|\mu - x|)$ with $\mu, x \in [0, 1, \ldots, 10]$. $TN(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2 truncated to the interval $(-1, 1)$ for the partial autocorrelations and $(0, 1000)$ for the standard deviation of the error term σ_ϵ respectively.

	True p	True q	α	α_w	α_b
Mean	2.41	2.49	0.25	0.34	0.05
Standard Deviation	1.7413	1.7436	0.13	0.17	0.03
Range	[0; 5]	[0; 5]	[0.02; 0.56]	[0.02; 0.69]	[0.00; 0.13]

Table 5.7: Acceptance rates for Cauchy distribution
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

ment with a share of correctly identified models of 0.35, the BIC outperforms RJMCMC in this setup with 40% of the models being correctly identified. The AIC and AICC perform slightly better than before, identifying the correct models in 20 and 22% of the cases respectively.

Criterion	Correct Identification
RJMCMC	0.35
AIC	0.20
AICC	0.22
BIC	0.40

Table 5.8: Share of correctly identified models for Cauchy distribution

Table 5.9 presents the familiar D_i statistic for the different approaches. While RJMCMC performs worse than the BIC with respect to the share of correctly identified models, its performance regarding the identification of the true impulse response shows a different picture. Taking this perspective, RJMCMC is competitive, with lowest values for means and medians for all estimates and horizons except for the median of D_i for the BIC at a horizon of 10 periods and the mean for the AIC at a horizon of 60 periods.

In terms of the variability of D_i across data sets, RJMCMC outperforms the other methods at horizons of 5 and 10 periods and is comparable to the other methods at longer horizons. In conclusion, the advantage of RJMCMC is substantially less pronounced than in the experiment using the mixture of normals, but RJMCMC is at the very least competitive. The results presented here reinforce the argument made in the foregoing regarding model uncertainty:

the ability of RJMCMC to account for model uncertainty may leave it at a disadvantage when the task is to pick one particular model as in the comparison of the model choices. The disadvantage of RJMCMC with regards to model selection compared vanishes completely when the impulse response mismatch is considered, which measures the ability of the methods to match the dynamics present in the data.

Horizon	5	10	20	40	60
RJMCMC Mean	0.217 [0.0895] (0.373)	0.38 [0.126] (0.854)	0.55 [0.154] (1.31)	0.836 [0.165] (2.32)	1.07 [0.165] (3.15)
RJMCMC Median	0.223 [0.0913] (0.387)	0.395 [0.118] (0.906)	0.577 [0.14] (1.4)	0.871 [0.152] (2.44)	1.1 [0.152] (3.31)
AIC	0.282 [0.112] (0.561)	0.465 [0.144] (1.01)	0.61 [0.182] (1.29)	0.839 [0.211] (1.89)	1.06 [0.249] (2.79)
AICC	0.283 [0.111] (0.568)	0.474 [0.144] (1.07)	0.642 [0.182] (1.47)	0.895 [0.196] (2.18)	1.14 [0.222] (3.13)
BIC	0.274 [0.101] (0.563)	0.461 [0.12] (1.11)	0.626 [0.156] (1.55)	0.883 [0.179] (2.33)	1.15 [0.179] (3.34)

Table 5.9: D_i for Cauchy distribution at different horizons
Mean, [Median], and (Standard Deviation)

5.6.3 Summary

RJMCMC performed better than the frequentist criteria in terms of matching the true impulse response for the disturbances following a mixture of two normal distributions. RJMCMC was competitive for the Cauchy distributed disturbance, where the BIC outperformed RJMCMC with respect to model selection, but not when considering the impulse response mismatch. In general, the BIC performed better than both AIC and AICC.

5.7 Sample Sizes

In this section, the influence of the sample size on the performance of the different approaches is investigated. The sample sizes n considered are 50, 150, 250, and 400. In order to best mimic reality, 100 data sets with 400 observations each were generated according to the method described in the foregoing. Additionally to sampling lag polynomial orders and parameters, the standard deviation of the disturbance was sampled from $U(0.1; 1.5)$. Then, the estimations were carried out using the first 50, 150, and so forth observations of this data set. This approach mimics reality in the sense that new data generated by the same data generating process arrives and the econometrician estimates the model with ever increasing amounts of data.[10]

5.7.1 Results

Table 5.10 presents the sampler settings employed in this experiment. Table 5.11 presents summary statistics of the acceptance rates achieved for the different sample sizes.

Object	Proposal
p	$DL(\mu, 2.2)$
q	$DL(\mu, 2.2)$
(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.03^2)$
σ_ϵ	$TN(\mu, 0.05^2)$

Table 5.10: Proposal distributions for sample size experiment
$DL(\mu, b)$ denotes the discretized Laplace distribution; $TN(\mu, \sigma^2)$ denotes the truncated normal distribution

The acceptance rates decrease with the sample size due to increased curvature in the likelihood, complicating the proposal of new states with similar posterior probability. Also

[10] It is very much likely, however, that e.g. the data generating process for GDP changes over time. Therefore, the degree of realism afforded by this approach may be limited.

n		α	α_w	α_b
50	Mean	0.41	0.56	0.09
	Standard Deviation	0.11	0.15	0.02
	Range	[0.09; 0.66]	[0.13; 0.85]	[0.02; 0.15]
150	Mean	0.28	0.38	0.06
	Standard Deviation	0.13	0.17	0.03
	Range	[0.05; 0.61]	[0.06; 0.76]	[0.01; 0.15]
250	Mean	0.21	0.29	0.04
	Standard Deviation	0.13	0.17	0.03
	Range	[0.03; 0.57]	[0.03; 0.71]	[0.00; 0.15]
400	Mean	0.16	0.22	0.03
	Standard Deviation	0.12	0.16	0.03
	Range	[0.01; 0.53]	[0.02; 0.65]	[0.00; 0.11]

Table 5.11: Acceptance rates for different sample sizes
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

here, an improvement of the performance of RJMCMC would be expected, had the sampler settings been tuned for each data set.

Table 5.12 contains the share of correctly identified models for the different sample sizes. Both BIC and RJMCMC deliver satisfactory performance even with a sample size of 150 where the share of correctly identified models is 0.34 and 0.33 respectively, while both AIC and AICC do not, irrespective of sample size. As expected, the performance of the methods improves with increasing sample size for RJMCMC and BIC. For both AIC and AICC the performance does not improve uniformly, with the AIC's performance leveling off at a sample size of 250 and the AICC's performance actually decreasing when increasing the sample size from 250 to 400.[11]

Turning to the analysis of the impulse response mismatch measures reported in Table 5.13, the changes in performance with increasing sample size mirrors the picture from the

[11] Again, this result is not surprising in light of the philosophy behind the Akaike criteria discussed above.

Criterion	50	150	250	400
RJMCMC	0.27	0.34	0.41	0.46
AIC	0.02	0.09	0.13	0.13
AICC	0.05	0.12	0.15	0.13
BIC	0.14	0.33	0.37	0.46

Table 5.12: Share of correctly identified models for different sample sizes

performance regarding model selection discussed above. While both BIC and RJMCMC show improved performance with increasing sample size across the board, both AIC and AICC do not, with the mean mismatch being roughly equal for sample sizes of 150 and 400 and a *degradation* in the mean for the intermediate sample size of 250. Considering the median of the mismatch measure, however, shows the uniform improvement one would expect also for these two criteria. These results are driven by some data sets for which AIC and AICC identified models which are far removed from the data generating process in terms of their dynamics.

RJMCMC again outperforms all other methods in this experiment. At all horizons and sample sizes both mean and median of the distribution of D_i are lower. Furthermore, RJMCMC estimates appear to be more efficient with regards to matching the impulse response as evident in the lower standard deviation of the measure across data sets. This effect is stronger the bigger the sample size. For example, considering a sample size of 400, the standard deviation for D_i calculated on the mean response from RJMCMC is 1.04, while the information criteria achieve a standard deviation which, at a value of 10.1 for all criteria, is one order of magnitude greater.

5.7.2 Summary

In conclusion, RJMCMC and BIC perform best in terms of model selection. AIC and AICC perform poorly, regardless of the sample size. When considering the ability of the methods to match the true impulse response functions, RJMCMC outperforms all methods by a substantial

Horizon	5	10	20	40	60
n = 50					
RJMCMC Mean	0.254 [0.0978] (0.518)	0.72 [0.193] (2.44)	1.39 [0.264] (5.97)	2.03 [0.334] (8.39)	2.32 [0.354] (8.85)
RJMCMC Median	0.252 [0.0979] (0.526)	0.723 [0.194] (2.62)	1.41 [0.257] (6.55)	2.06 [0.318] (9.19)	2.36 [0.333] (9.76)
AIC	0.447 [0.205] (0.658)	1.25 [0.459] (2.81)	2.46 [0.616] (7.43)	3.7 [0.7] (10.8)	4.31 [0.728] (12)
AICC	0.416 [0.196] (0.649)	1.23 [0.404] (2.86)	2.52 [0.507] (7.68)	3.87 [0.579] (11.6)	4.46 [0.606] (12.7)
BIC	0.384 [0.162] (0.715)	1.14 [0.339] (2.9)	2.29 [0.398] (7.55)	3.51 [0.477] (11.3)	4.18 [0.574] (12.7)
n = 150					
RJMCMC Mean	0.0985 [0.0286] (0.212)	0.281 [0.058] (1.02)	0.589 [0.0729] (2.55)	0.894 [0.0871] (3.77)	1.04 [0.0873] (4.08)
RJMCMC Median	0.0973 [0.0276] (0.21)	0.279 [0.0504] (1.05)	0.604 [0.0725] (2.79)	0.951 [0.0812] (4.46)	1.11 [0.0843] (4.86)
AIC	0.134 [0.0501] (0.359)	0.385 [0.109] (1.42)	0.752 [0.16] (2.43)	1.37 [0.199] (5.19)	1.63 [0.212] (5.62)
AICC	0.138 [0.047] (0.366)	0.388 [0.0873] (1.42)	0.766 [0.156] (2.46)	1.39 [0.198] (5.21)	1.65 [0.212] (5.64)
BIC	0.133 [0.0391] (0.377)	0.393 [0.076] (1.47)	0.704 [0.107] (2.4)	1.05 [0.138] (3.43)	1.29 [0.158] (4.32)
n = 250					
RJMCMC Mean	0.0418 [0.0162] (0.0806)	0.115 [0.0312] (0.374)	0.231 [0.0366] (0.821)	0.354 [0.0467] (1.02)	0.448 [0.0502] (1.24)
RJMCMC Median	0.0409 [0.0165] (0.0757)	0.114 [0.0309] (0.337)	0.235 [0.0373] (0.783)	0.372 [0.044] (1.07)	0.474 [0.0459] (1.35)
AIC	0.0743 [0.0285] (0.179)	0.229 [0.072] (0.836)	0.6 [0.101] (2.99)	1.52 [0.126] (7.96)	2.09 [0.138] (11.8)
AICC	0.072 [0.0301] (0.179)	0.225 [0.0642] (0.837)	0.585 [0.0903] (2.99)	1.51 [0.114] (7.96)	2.07 [0.129] (11.8)
BIC	0.049 [0.028] (0.0923)	0.122 [0.0404] (0.24)	0.315 [0.0553] (0.967)	1.07 [0.0615] (6.54)	1.51 [0.0788] (9.61)
n = 400					
RJMCMC Mean	0.027 [0.00832] (0.0506)	0.0736 [0.017] (0.226)	0.159 [0.0247] (0.562)	0.254 [0.0313] (0.938)	0.306 [0.039] (1.04)
RJMCMC Median	0.027 [0.00777] (0.0494)	0.0736 [0.019] (0.215)	0.163 [0.0256] (0.601)	0.27 [0.0315] (1.14)	0.326 [0.0375] (1.31)
AIC	0.0493 [0.0195] (0.102)	0.151 [0.0419] (0.543)	0.482 [0.0589] (2.37)	1.29 [0.0718] (7.74)	1.63 [0.0822] (10.1)
AICC	0.0493 [0.0195] (0.102)	0.151 [0.0419] (0.543)	0.482 [0.0589] (2.37)	1.29 [0.0718] (7.74)	1.63 [0.0822] (10.1)
BIC	0.0394 [0.00861] (0.0979)	0.127 [0.0187] (0.54)	0.437 [0.031] (2.36)	1.22 [0.043] (7.73)	1.56 [0.0479] (10.1)

Table 5.13: D_i for different sample sizes across data sets
Mean, [Median], and (Standard Deviation)

margin, especially as the sample size increases. Not only are mean and median mismatches lower, irrespective of the choice of estimate from RJMCMC,[12] the estimates of D_i for RJMCMC also exhibit a substantially smaller standard deviation across data sets, especially at larger sample sizes and longer horizons, suggesting that RJMCMC estimates of the dynamics are more efficient.

5.8 Trend Break

This section is concerned with the effects of a break in a linear trend in the data which is ignored by the econometrician. Data for this experiment is based on the estimates for US real GDP from Morley, Nelson, and Zivot (2003) following Perron and Wada (2009). The data generating process is given by

$$\begin{aligned} y_t = a + \mu t + d\mathbb{1}(t > T_b)(t - T_b) + c_t \\ c_t = \Phi_1 c_{t-1} + \Phi_2 c_{t-2} + \epsilon_t \\ \epsilon_t \sim i.i.d.\, N(0, \sigma_\epsilon^2) \end{aligned}$$

with $a = 724.18, \mu = 0.95, \Phi_1 = 1.28, \Phi_2 = -0.38, d = -0.29, \sigma_\epsilon = 0.94$. $\mathbb{1}(t > T_b)$ is the indicator function equal to one if $t > T_b$. 100 synthetic data sets were generated with a sample size of 200 and the break in the trend in the middle of the sample, i.e. $T_b = 100$, as in Perron and Wada (2009). With these nonstationary data, the estimation procedure was run using linear detrending and first differencing for each data set. RJMCMC estimates are again based on every 30th sample after the burn in, except for the binary model choice for which the full sample from the posterior is used.

[12] That is, based on mean or median responses.

5.8.1 Persistence and Stationarity

The results for real GDP in chapter 4 may suggest that the time series studied are difference stationary based on the estimates of the persistence measure employed there, consistent with the findings of Campbell and Mankiw (1987) and Campbell and Mankiw (1989). This angle is not further investigated in chapter 4, a decision supported by the results presented here which suggest that this measure may not be able to identify trend stationarity in this setup.[13]

The following exposition of the persistence measure is taken from chapter 4. Let $C(L)$ denote the infinite order polynomial in the lag operator given by the infinite moving average representation of a stationary ARMA(p,q) model and let $C_n(1)$ be the sum of the first n coefficients:

$$P(L)y_t = Q(L)\epsilon_t \tag{5.4}$$

$$y_t = \frac{Q(L)}{P(L)}\epsilon_t = C(L)\epsilon_t = (1 + C_1 L + C_2 L^2 + \ldots)\epsilon_t \tag{5.5}$$

$$C_n(1) = \sum_{i=1}^{n} 1 + C_i \tag{5.6}$$

$C_n(1)$ thus gives the *cumulated* response to a shock up to horizon n.

What information does this statistic convey? Consider first a model in which the y_t are first-differenced log GDP per capita data points. In this setup, C_i gives the effect of a disturbance on the *growth rate* occurring at time t on the *growth rate* at time $t + i$. The cumulative effect on the *level* of GDP at time $t + n$ is then given by $C_n(1)$. $C_n(1)$ is thus the change in one's forecast for the level of GDP at time $t + n$ after observing a unit shock in t. For a random walk, $C_n(1) = 1 \forall n$ would hold, for example, while if the series were trend-stationary $C_n(1)$ would converge to zero with increasing n as the effect of the shock on the level of GDP vanishes with trend-reversion (see Campbell and Mankiw (1987) for further discussion).

In the following, the persistence measure estimates from the non-stationary data will be

[13] It is by far not unlikely, that a permanent shift in the growth rate of GDP is present in real world data.

analyzed in order to answer the question whether the method would recognize the trend stationarity of the synthetic data by providing estimates of $C_n(1)$ for first differenced data which are low and converge to zero as the horizon increases. The persistence measure estimates from RJMCMC were based on every 30th draw from the posterior after burn-in to keep computations manageable.

5.8.2 First Differencing

For this experiment, the data generated according to the data generating process described above was differenced once and demeaned as in chapter 4. In essence, the data was thus mistakenly assumed to be difference stationary. Table 5.14 presents the sampler settings for this experiment and Table 5.15 the acceptance rates achieved.

Object	Proposal
p	$DL(\mu, 2.2)$
q	$DL(\mu, 2.2)$
(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.12^2)$
σ_ϵ	$TN(\mu, 0.05^2)$

Table 5.14: Proposal distributions for first differences
$DL(\mu, b)$ denotes the discretized Laplace distribution; $TN(\mu, \sigma^2)$ denotes the truncated normal distribution

	α	α_w	α_b
Mean	0.25	0.32	0.08
Standard Deviation	0.05	0.07	0.01
Range	[0.05; 0.42]	[0.05; 0.52]	[0.03; 0.14]

Table 5.15: Acceptance rates for first differences
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

Since the purpose of this experiment is to evaluate whether the estimates of $C_n(1)$ point towards the trend stationary nature of the data, I turn directly to the presentation of the corresponding results.

Table 5.16 presents the following estimates for $C_n(1)$ at different horizons. The left column indicates the source of the estimate for each data set. The mean (median) estimate is based on the mean (median) of $C_n(1)$ at each horizon across samples from the posterior. For each of the horizons, the table then reports the mean and (standard deviation) across data sets, as well as the range of the estimates in the second row.

Horizon	5	10	20	40	60
RJMCMC	$1.39\ (0.128)$	$1.38\ (0.137)$	$1.38\ (0.138)$	$1.38\ (0.138)$	$1.37\ (0.138)$
Mean	$[1.03; 1.75]$	$[0.939; 1.75]$	$[0.912; 1.75]$	$[0.906; 1.75]$	$[0.905; 1.75]$
RJMCMC	$1.39\ (0.125)$	$1.38\ (0.137)$	$1.38\ (0.139)$	$1.38\ (0.139)$	$1.38\ (0.139)$
Median	$[0.98; 1.74]$	$[0.878; 1.74]$	$[0.837; 1.74]$	$[0.831; 1.74]$	$[0.831; 1.74]$
AIC	$1.25\ (0.24)$	$1.12\ (0.325)$	$1.15\ (0.306)$	$1.13\ (0.313)$	$1.13\ (0.319)$
	$[0.812; 2.08]$	$[0.385; 1.8]$	$[0.524; 1.92]$	$[0.257; 1.85]$	$[0.0999; 1.86]$
AICC	$1.25\ (0.239)$	$1.12\ (0.32)$	$1.14\ (0.309)$	$1.12\ (0.311)$	$1.13\ (0.315)$
	$[0.812; 2.08]$	$[0.385; 1.8]$	$[0.524; 1.92]$	$[0.257; 1.85]$	$[0.0999; 1.86]$
BIC	$1.43\ (0.176)$	$1.42\ (0.215)$	$1.42\ (0.214)$	$1.42\ (0.214)$	$1.42\ (0.214)$
	$[0.847; 1.92]$	$[0.612; 1.95]$	$[0.605; 1.95]$	$[0.606; 1.95]$	$[0.606; 1.95]$

Table 5.16: $C_n(1)$ for first differences at different horizons
Mean, (Standard Deviation), and [Range]

Clearly, the estimates do not converge to zero as the horizon increases and are similar to those in chapter 4. This supports the results from Perron (1989), who shows that neglecting a one-time break in a non-stochastic time trend leads to inability to reject the unit root hypothesis against trend stationary alternatives in standard tests, even asymptotically.

5.8.3 Linear Detrending

In this experiment, the data were taken to be stationary about a linear trend, again without accounting for the break in the trend in the middle of the sample. The data was detrended using OLS detrending and the resulting cyclical component was used for estimation.

Table 5.17 presents the sampler settings employed here. Table 5.18 reports the familiar summary statistics for the acceptance rates. Table 5.19 presents the share of correctly identified models for the cyclical component.

Object	Proposal
p	$DL(\mu, 2.2)$
q	$DL(\mu, 2.2)$
(Inverse) Partial Autocorrelation Between	$TN(\mu, 0.05^2)$
(Inverse) Partial Autocorrelation Within	$TN(\mu, 0.05^2)$
σ_ϵ	$TN(\mu, 0.05^2)$

Table 5.17: Proposal distributions for linear trend
$DL(\mu, b)$ denotes the discretized Laplace distribution; $TN(\mu, \sigma^2)$ denotes the truncated normal distribution

In this setup, the BIC identifies the correct model in a surprisingly large number of cases and the AIC as well as the AICC perform substantially worse. Both AIC and AICC tend to choose less parsimonious models than either RJMCMC or BIC. This phenomenon has also been observed in chapter 4 and the reasons for this phenomenon have been discussed above.

	α	α_w	α_b
Mean	0.20	0.27	0.05
Standard Deviation	0.02	0.03	0.01
Range	[0.13; 0.25]	[0.18; 0.36]	[0.04; 0.06]

Table 5.18: Acceptance rates for linear trend
α denotes the overall acceptance rate; α_w and α_b denote acceptance rates for within and between model moves respectively

Criterion	Correct Identification
RJMCMC	0.54
AIC	0.11
AICC	0.16
BIC	0.72

Table 5.19: Share of correctly identified models for linear trend

Here, it is very pronounced: while the average sum of lags $p+q$ chosen by BIC and RJMCMC is 2.1 and 2.24 respectively, this measure is 6.88 (6.33) for the AIC (AICC).

Horizon	5	10	20	40	60
RJMCMC Mean	$0.514\,[0.42]$ (0.322)	$2.61\,[2.48]$ (1.05)	$7.55\,[7.45]$ (2.63)	$12.6\,[12.2]$ (4.84)	$14.5\,[13.5]$ (5.98)
RJMCMC Median	$0.481\,[0.385]$ (0.307)	$2.49\,[2.34]$ (0.998)	$7.18\,[7.01]$ (2.57)	$11.5\,[10.9]$ (4.71)	$12.7\,[11.7]$ (5.62)
AIC	$0.475\,[0.369]$ (0.389)	$2.37\,[2.1]$ (1.48)	$7.33\,[6.2]$ (4.25)	$13.2\,[10.5]$ (9.49)	$15.9\,[11.8]$ (13.8)
AICC	$0.478\,[0.375]$ (0.381)	$2.4\,[2.16]$ (1.43)	$7.34\,[6.5]$ (4.1)	$13\,[10.8]$ (9.18)	$15.6\,[11.9]$ (13.3)
BIC	$0.486\,[0.4]$ (0.352)	$2.39\,[2.15]$ (1.15)	$6.52\,[6.22]$ (2.86)	$10\,[9.3]$ (5.05)	$10.9\,[9.86]$ (5.9)

Table 5.20: D_i for linear trend at different horizons
Mean, [Median], and (Standard Deviation)

Even though the acceptance rates were better for this experiment, RJMCMC is competitive merely with both AIC and AICC when considering the impulse response mismatch measure. The standard deviation of the mismatch measure across data sets is lower for RJMCMC and the BIC and the BIC performs best in terms of matching the true impulse response, followed by RJMCMC.

5.8.4 Summary

To summarize, RJMCMC remained competitive with the other methods when considering the OLS detrended case. Again, the BIC performed best among the information criteria and outperformed RJMCMC under linear detrending. This section has also shown that tests of difference vs trend stationary hypotheses may not be able to identify trend stationarity if the deterministic trend features a break in the trend function that is not modeled.

5.9 Conclusion

This study has investigated the performance of RJMCMC in comparison to frequentist approaches to the estimation of ARMA(p,q) models in several settings.

The first two experiments considered here employed data generated from a mixture of two normals, as well as a truncated Cauchy distribution. In these experiments, the standard Gaussian likelihood function was thus misspecified. While AIC and AICC performed significantly worse, RJMCMC was comparable to the BIC in terms of model selection with RJMCMC performing slightly better in the case of normal mixtures and the BIC outperforming RJMCMC using Cauchy distributed disturbances. Regarding the mismatch of the estimated and true impulse response functions, however, RJMCMC outperformed all methods in both experiments.

When considering the influence of sample size on the quality of the results, both RJMCMC and BIC delivered similar performance. The results for both AIC and AICC were significantly worse with respect to model selection, where both RJMCMC and BIC deliver satisfactory results starting with a sample size of only 150. Even with a sample size of 50, RJMCMC identified the correct model in 27% of cases, followed by the BIC with 14%. AIC and AICC identified the correct model only in 2% and 5% of cases respectively. The improvement of the performance with increasing sample size was monotonous for RJMCMC and BIC, but neither for the AIC, nor the AICC. For the latter two, performance did not improve after a sample size of 250 with the performance actually degrading for the AICC when increasing the sample size to 400. Regarding the impulse response mismatch measure, RJMCMC outperformed all other methods across all measures considered with the BIC coming in second place. Most importantly, RJMCMC estimates of the dynamics of the data generating process seem to be more efficient compared to the other approaches, especially with larger sample sizes and longer horizons. For example, the standard deviation of D_i across data sets for RJMCMC was one tenth that of the value for the other methods at a sample size of 400 and a horizon of 60 periods.

The last experiment concerned the question of the effects of a disregarded one-time break in a deterministic trend. Two popular detrending methods were applied to data generated with a break in the deterministic trend in the middle of the sample. When the data was OLS-detrended, RJMCMC and BIC delivered similar results. With respect to identification of one particular model, the BIC correctly identified the model in 72% of cases, compared to 54% for RJMCMC, and 11% (16%) for the AIC (AICC). In terms of the ability to match the true impulse response, RJMCMC came in second place after the BIC.

In the case of detrending using first differencing, the estimates for all four approaches for the persistence measure employed in chapter 4 and Campbell and Mankiw (1987), among others, were analyzed. The results suggest, that inference regarding the question whether economic time series are trend or difference stationary should not be based on estimates of the persistence measure when a break in a deterministic time trend is present but not modeled correctly since the estimates do not converge to zero with increasing horizon. This is in line with the results in Perron (1989).

To conclude, in all the experiments RJMCMC delivered performance which was at the very least competitive compared to the other approaches. In general, RJMCMC and BIC performed best, with AIC and AICC performing worse, at times substantially, both in terms of model selection as well as the mismatch between the true and estimated impulse response functions. Furthermore, inference regarding hypotheses of trend and difference stationarity based on the persistence measure of Campbell and Mankiw (1987) should be avoided.

Bibliography

Akaike, H. (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Alejandro Justiniano, G. E. P., and A. Tambalotti (2008): "Investment Shocks and Business Cycles," Discussion paper, Federal Reserve Bank of Chicago.

Alessi, L., M. Barigozzi, and M. Capasso (2011): "Non-Fundamentalness in Structural Econometric Models: A Review," *International Statistical Review*, 79(1), 16–47.

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172.

Anderson, G. S. (2010): "A Reliable and Computationally Efficient Algorithm for Imposing the Saddle Point Property in Dynamic Models," *Journal of Economic Dynamics and Control*, 34(3), 472–489.

Barndorff-Nielsen, O., and G. Schou (1973): "On the Parametrization of Autoregressive Models by Partial Autocorrelations," *Journal of Multivariate Analysis*, 3, 408–419.

Barsky, R. B., and E. R. Sims (2011): "News shocks and business cycles," *Journal of Monetary Economics*, 58(3), 273 – 289.

BAXTER, M., AND R. G. KING (1999): "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series," *The Review of Economics and Statistics*, 81(4), 575–593.

BAYARRI, M., AND J. BERGER (2004): "The Interplay of Bayesian and Frequentist Analysis," *Statistical Science*, 19(1), 58–80.

BEAUDRY, P., AND F. PORTIER (2005): "The news view of economic fluctuations: Evidence from aggregate Japanese data and sectoral US data," *Journal of the Japanese and International Economies*, 19(4), 635–652.

BLANCHARD, O. J., AND S. FISCHER (1989): *Lectures on Macroeconomics*. MIT Press, Cambridge, MA.

BROOKS, S. P., AND R. EHLERS (2004): "Bayesian Analysis of Order Uncertainty in ARIMA Models," Discussion paper, Federal University of Parana, Brazil; University of Cambridge, UK.

BROOKS, S. P., AND A. GELMAN (1998): "General Methods for Monitoring Convergence of Iterative Simulations," *Journal of Computational and Graphical Statistics*, 7(4), 434–455.

BROOKS, S. P., P. GIUDICI, AND G. O. ROBERTS (2003): "Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1), 3–39.

CAMPBELL, J., AND G. MANKIW (1989): "International Evidence on the Persistence of Economic Fluctuations," *Journal of Monetary Economics*, 23, 319–333.

CAMPBELL, J. Y., AND N. G. MANKIW (1987): "Are Output Fluctuations Transitory?," *The Quarterly Journal of Economics*, 102(4), 857–80.

CANOVA, F. (1998): "Detrending and business cycle facts," *Journal of Monetary Economics*, 41, 475–512.

Cappè, O., C. P. Robert, and T. Rydèn (2003): "Reversible jump, birth-and-death and more general continuous time Markov chain Monte Carlo samplers," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(3), 679–700.

Chari, V., P. Kehoe, and E. McGrattan (2008): "Are structural VARs with long-run restrictions useful in developing business cycle theory?," *Journal of Monetary Economics*, 55, 1337–1352.

Cheung, Y.-W. (1994): "Aggregate output dynamics in the twentieth century," *Economics Letters*, 45, 15–22.

Cheung, Y.-W., and K. S. Lai (1992): "International evidence on output persistence from postwar data," *Economic Letters*, 38, 435–442.

Chib, S., and E. Greenberg (1995): "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49(4), 327–335.

Chow, G. C. (1981): "A Comparison of the Information and Posterior Probability Criteria for Model Selection," *Journal of Econometrics*, 16, 21–33.

Christiano, L. J., and M. Eichenbaum (1990): "Unit roots in real GNP: Do we know, and do we care?," *Carnegie-Rochester Conference Series on Public Policy*, 32(0), 7 – 61.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Clark, P. K. (1989): "Trend Reversion in Real Output and Unemployment," *Journal of Econometrics*, 40, 15–32.

Cochrane, J. H. (1988): "How Big Is the Random Walk in GNP?," *Journal of Political Economy*, 96(5), 893–920.

Cogley, T., and J. M. Nason (1995a): "Effects of the Hodrick-Prescott filter on trend and difference stationary time series - Implications for business cycle research," *Journal of Economic Dynamics and Control*, 19, 253–278.

Cogley, T., and J. M. Nason (1995b): "Output Dynamics in Real-Business-Cycle Models," *American Economic Review*, 85(3), 492–511.

Croce, M. M. (2014): "Long-run productivity risk: A new hope for production-based asset pricing?," *Journal of Monetary Economics*, 66, 13–31.

Cúrdia, V., and R. Reis (2010): "Correlated Disturbances and U.S. Business Cycles," NBER Working Papers 15774, National Bureau of Economic Research, Inc.

DeJong, D. N., and C. H. Whiteman (1991): "Reconsiderung 'Trends and random walks in macroeconomic time series'," *Journal of Monetary Economics*, 28, 221–254.

Del Negro, M., and F. Schorfheide (2006): "How Good Is What You've Got? DSGE-VAR As a Toolkit for Evaluating DSGE Models," *Economic Review*, 91(2), 21–37.

——— (2009): "Monetary Policy Analysis with Potentially Misspecified Models," *American Economic Review*, 99(4), 1415–50.

——— (2011): "Bayesian Macroeconometrics," in *Handbook of Bayesian Econometrics*, Chapters, chap. 7, p. 293Ű389. Oxford University Press.

Diebold, F. X., and G. D. Rudebusch (1989): "Long memory and persistence in aggregate output," *Journal of Monetary Economics*, 24(2), 189–209.

Durlauf, S. N., D. Romer, and C. A. Sims (1989): "Output Persistence, Economic Structure, and the Choice of Stabilization Policy," *Brookings Papers on Economic Activity*, 2, 66–136.

EHLERS, R., AND S. P. BROOKS (2002): "Efficient Construction of Reversible Jump MCMC Proposals for Autoregressive Time Series Models," Discussion paper, Statistical Laboratory Research Report, University of Cambridge, UK.

EHLERS, R., AND S. P. BROOKS (2008): "Adaptive Proposal Construction for Reversible Jump MCMC," *Scandinavian Journal of Statistics*, 35(4), 677–690.

FAHIMAH AL-AWADHI, M. H., AND C. JENNISON (2004): "Improving the acceptance rate of reversible jump MCMC proposals," *Statistics & Probability Letters*, 69, 189–198.

FAN, Y., AND S. SISSON (2011): *Handbook of Markov Chain Monte Carlo*chap. 3, pp. 67–87. CRC Press.

FATAS, A. (2000a): "Do Business Cycles Cast Long Shadows? Short-Run Persistence and Economic Growth," *Journal of Economic Growth*, 5, 147–162.

——— (2000b): "Endogenous growth and stochastic trends," *Journal of Monetary Economics*, 45, 107–128.

FERNÁNDEZ-VILLAVERDE, J., P. GUERRÓN-QUINTANA, AND J. F. RUBIO-RAMÍREZ (2010): "The New Macroeconometrics: A Bayesian Approach," in *Handbook of Applied Bayesian Analysis*, ed. by A. O'Hagan, and M. West, chap. 15. Oxford University Press.

FERNÁNDEZ-VILLAVERDE, J., AND J. F. RUBIO-RAMÍREZ (2004): "Comparing dynamic equilibrium models to data: a Bayesian approach," *Journal of Econometrics*, 123(1), 153 – 187.

FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMIREZ, T. J. SARGENT, AND M. W. WATSON (2007): "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97(3), 1021–1026.

FRANCIS, N., AND V. A. RAMEY (2005): "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited," *Journal of Monetary Economics*, 52(8), 1379–1399.

FREEDMAN, D. (1997): *Topics in the Foundation of Statistics*. Springer Science + Business Media B.V.

GALI, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 89(1), 249–271.

GELFAND, A. E., AND A. F. M. SMITH (1990): "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85(410), pp. 398–409.

GELMAN, A. (2008): "Objections to Bayesian Statistics," *Bayesian Analysis*, 3(3), 445–450.

GELMAN, A., G. ROBERTS, AND W. GILKS (1996): "Efficient Metropolis jumping rules," in *Bayesian Statistics*, ed. by J. M. Bernado, et al., vol. 5, p. 599. OUP.

GEWEKE, J. (1998): "Using simulation methods for Bayesian econometric models: inference, development, and communication," Staff Report 249, Federal Reserve Bank of Minneapolis.

GREEN, P. J. (1995): "Reversible Jump Markov Chain Monte Carlo," *Biometrika*, 82, 711–732.

GUERRÓN-QUINTANA, P., AND J. M. NASON (2013): "Bayesian estimation of DSGE models," in *Handbook of Research Methods and Applications in Empirical Macroeconomics*, Chapters, chap. 21, pp. 486–512. Edward Elgar.

HAMILTON, J. (1994): *Time series analysis*. Princeton Univ. Press, Princeton, NJ.

HANSEN, G. D. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16(3), 309–327.

HASTINGS, W. (1970): "Monte Carlo Sampling Methods using Markov Chains and Their Applications," *Biometrika*, 57, 97–109.

HERBST, E., AND F. SCHORFHEIDE (2015): "Bayesian Inference for DSGE Models," Mimeo.

HODRICK, R. J., AND E. C. PRESCOTT (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29(1), 1–16.

HOETING, J. A., D. MADIGAN, A. E. RAFTERY, AND C. T. VOLINSKY (1999): "Bayesian Model Averaging: A Tutorial," *Statistical Science*, 14(4), 382–417.

HURVICH, C. M., AND C.-L. TSAI (1989): "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76(2), 297–307.

JONES, C. I. (1995): "Time Series Tests of Endogenous Growth Models," *The Quarterly Journal of Economics*, 110(2), 495–525.

JONES, M. (1987): "Randomly Choosing Parameters from the Stationarity and Invertibility Region of Autoregressive-Moving Average Models," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 36, 134–138.

KASS, R. E., AND A. E. RAFTERY (1995): "Bayes factors," *Journal of the American Statistical Association*, 90, 773–795.

KING, R. G., AND S. T. REBELO (1993): "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control*, 17, 207–231.

KLEIN, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

KOOP, G. (1991): "Intertemporal Properties of Real Output: A Bayesian Analysis," *Journal of Business & Economic Statistics*, 9(3), 253–265.

KOOP, G., R. LEON-GONZALEZ, AND R. STRACHAN (2012): "Bayesian model averaging in the instrumental variable regression model," *Journal of Econometrics*, 171(2), 237 – 250, Bayesian Models, Methods and Applications.

KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT, AND Y. SHIN (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?," *Journal of Econometrics*, 54(1-3), 159–178.

KYDLAND, F. E., AND E. C. PRESCOTT (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–70.

LAN, H., AND A. MEYER-GOHDE (2012): "Existence and Uniqueness of Perturbation Solutions to DSGE Models," SFB 649 Discussion Papers 2012-015.

LEAMER, E. E. (1978): *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. Wiley.

LIPPI, M., AND L. REICHLIN (1994): "VAR analysis, nonfundamental representations, blaschke matrices," *Journal of Econometrics*, 63(1), 307–325.

MADIGAN, D., AND A. E. RAFTERY (1994): "Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window," *Journal of the American Statistical Association*, 89(428), 1535–1546.

METROPOLIS, N., A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, AND E. TELLER (1953): "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, 21(6), 1087–1092.

MEYER-GOHDE, A. (2010): "Linear Rational-Expectations Models with Lagged Expectations: A Synthetic Method," *Journal of Economic Dynamics and Control*, 34(5), 984–1002.

MONAHAN, J. (1984): "A note on enforcing stationarity in autoregressive-moving average models," *Biometrika*, 71, 403–404.

Morley, J. C., C. R. Nelson, and E. Zivot (2003): "Why are the Beveridge-Nelson and Unobserved-Components Decomposition of GDP so Different?," *The Review of Economics and Statistics*, 135(2).

Nelson, C., and C. Plosser (1982): "Trends and Random Walk in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139–162.

Perron, P. (1989): "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57(6), 1361–1401.

——— (1993): "The HUMP-Shaped Behavior of Macroeconomic Fluctuations," *Empirical Economics*, 18, 707–727.

Perron, P., and T. Wada (2009): "Let's take a break: Trends and cycles in US real GDP," *Journal of Monetary Economics*, 56, 749–765.

Peskun, P. (1973): "Optimum Monte-Carlo sampling using Markov chains," *Biometrika*, 60, 607–612.

Philippe, A. (2006): "Bayesian analysis of autoregressive moving average processes with unknown orders," *Computational Statistics & Data Analysis*, 51, 1904–1923.

Quenouille, M. (1957): *The Analysis of Multiple Time-Series*. Hafner Publishing Company.

Raftery, A. E., D. Madigan, and J. A. Hoeting (1997): "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association*, 92(437), 179–191.

Ragacs, C., and M. Zagler (2002): "Persistence of Shocks to Output in Austria and Theories of Economic Growth," *Empirica*, 29, 305–317.

Ravenna, F. (2007): "Vector autoregressions and reduced form representations of DSGE models," *Journal of Monetary Economics*, 54, 2048–2064.

SARGENT, T. J. (1987): *Macroeconomic Theory*. Academic Press, San Diego, CA, 2nd edn.

SCHMITT-GROHÉ, S., AND M. URIBE (2010): "Evaluating the Sample Likelihood of Linearized DSGE Models without the Use of the Kalman Filter," *Economics Letters*, 109(3), 142–143.

SCHWARZ, G. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics*, 6(2), 461–464.

SIMS, C. A. (1998): "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, 49(1), 317–356.

——— (2001): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20.

——— (2012): "Statistical Modeling of Monetary Policy and Its Effects," *American Economic Review*, 102(4), 1187–1205.

SISSON, S. A. (2005): "Transdimensional Markov Chains: A Decade of Progress and Future Perspectives," *Journal of the American Statistical Association*, 100(471), 1077–1089.

SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

STEPHENS, M. (2000): "Bayesian analysis of mixture models with an unknown number of componentsŮan alternative to reversible jump methods," *The Annals of Statistics*, 28(1), 40–74.

STOCK, J. H., AND M. W. WATSON (1988): "Variable Trends in Economic Time Series," *The Journal of Economic Perspectives*, 2(3), pp. 147–174.

STRACHAN, R. W., AND H. K. VAN DIJK (2013): "Evidence On Features Of A Dsge Business Cycle Model From Bayesian Model Averaging," *International Economic Review*, 54(1), 385–402.

SUGIURA, N. (1978): "Further analysts of the data by akaike' s information criterion and the finite corrections," *Communications in Statistics - Theory and Methods*, 7(1), 13–26.

TIERNEY, L. (1994): "Markov Chains for Exploring Posterior Distributions," *The Annals of Statistics*, 22(4), 1701–1762.

——— (1998): "A note on Metropolis-Hastings kernels for general state spaces," *The Annals of Applied Probability*, 8(1), 1–9.

TROUGHTON, P., AND S. J. GODSILL (1998): "A Reversible Jump Sampler for Autoregressive Time Series, Employing Full Conditionals to Achieve Efficient Model Space Moves," pp. 2257–2260.

UHLIG, H. (1999): "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott, chap. 3, pp. 30–61. Oxford University Press.

——— (2004): "Do Technology Shocks Lead to a Fall in Total Hours Worked?," *Journal of the European Economic Association*, 2(2-3), 361–371.

WAAGEPETERSEN, R., AND D. SORENSEN (2001): "A Tutorial on Reversible Jump MCMC with a View toward Applications in QTL-mapping," *International Statistical Review*, 69, 49–61.

WATSON, M. (1993): "Measures of Fit for Calibrated Models," *Journal of Political Economy*, 101, 1011–1041.

WATSON, M. W. (1986): "Univariate Detrending Methods With Stochastic Trends," *Journal of Monetary Economics*, 18, 49–75.

WOLTERS, M. H. (2015): "Evaluating Point and Density Forecasts of DSGE Models," *Journal of Applied Econometrics*, 30(1), 74–96.